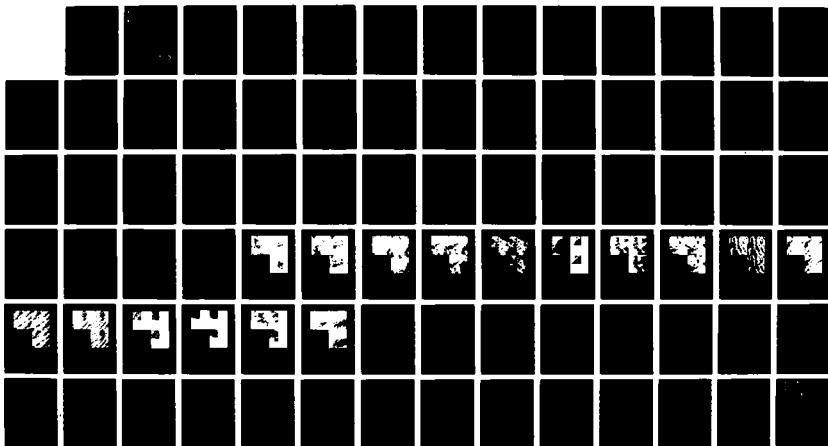MODELING CLOUDY AND CLEAR INTERVAL LENGTH PROBABILITIES
USING SPACE SHUTTLE IMAGERY(U) AIR FORCE INST OF TECH
WRIGHT-PATTERSON AFB OH   G F HOWARD 1987
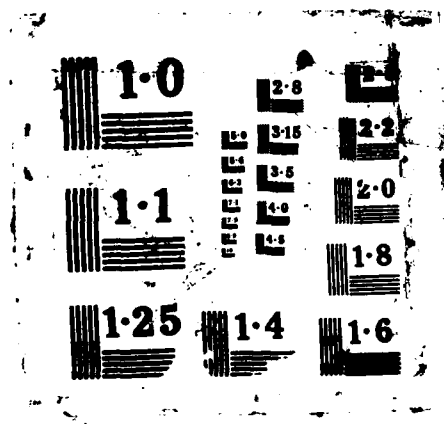
AFIT/CI/NR-87-36T

F/G 4/2   NL

AD-A185 290

DTIC FILE COPY

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFIT/CI/NR 87-36T	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Modeling Cloudy and Clear Interval Length Probabilities Using Space Shuttle Imagery		5. TYPE OF REPORT & PERIOD COVERED THESIS/DISSERTATION
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) George Franklin Howard III		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS AFIT STUDENT AT: Colorado State University		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS AFIT/NR WPAFB OH 45433-6583		12. REPORT DATE 1987
		13. NUMBER OF PAGES 71
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES
APPROVED FOR PUBLIC RELEASE: IAW AFR 190-1

LYNN E. WOLAVER 17 Aug 87
Dean for Research and
Professional Development
AFIT/NR

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

DTIC
ELECTE
OCT 2 6 1987
S   D

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
ATTACHED

DD FORM 1473   EDITION OF 1 NOV 65 IS OBSOLETE
1 JAN 73

THESIS

MODELING CLOUDY AND CLEAR INTERVAL LENGTH
PROBABILITIES USING SPACE SHUTTLE IMAGERY

Submitted by
George Franklin Howard III
Department of Atmospheric Science

COLORADO STATE UNIVERSITY

<u>May 21</u> 19<u>87</u>

WE HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER OUR SUPER-
VISION

BY_____George Franklin Howard, III_____

ENTITLED__Modeling Cloudy and Clear Interval Length Probabilities__

_____Using Space Shuttle Imagery_____

BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF

_____Masters of Science_____

<u>COMMITTEE ON GRADUATE WORK</u>

ADVISER

DEPARTMENT HEAD

ABSTRACT

Interval length probabilities provide an alternative to other characterizations of cloudy and clear regions as viewed from atop the atmosphere. This work attempts to accurately model these probabilities using very high resolution space shuttle orbiter images. Probabilities extracted from these images are compared with three model representations. Metric and congruent statistical methods based on absolute deviations are used to determine model goodness-of-fit. An exponential model is shown to exhibit the least error of the three. Further examination shows that the parameters used to fit the exponential model to observed probabilities can be obtained from the cloud field itself (in the form of mean cloudy and mean clear interval lengths). These mean values are determined for image fractions as small as 1/32 and used to predict probabilities for the entire image.

George Franklin Howard III
Department of Atmospheric Science
Colorado State University
Fort Collins, Colorado 80523
Summer, 1987

ACKNOWLEDGEMENTS

The author wishes to thank Dr. Thomas H. Vonder Haar for his support and advice throughout this study. Thanks also to Dr. Stephen K. Cox and Dr. Paul W. Mielke, Jr. for their service as committee members, their helpful discussions, and careful review of this thesis.

Particular thanks are extended to Dr. J. William Snow of the Air Force Geophysics Laboratory for many hours of critical discussion, provision of the data used in this study, and service as a committee member; and to Lieutenant Colonel Edward M. Tomlinson for his role in initiating and coordinating the CLOUDS experiment.

The author also wishes to thank Captain Francis P. Kelly for indoctrination into the world of metric and congruent statistical methods, and Mr. Chi-Fan Shih for considerable and excellent programming support, and Mrs. Loretta Wilson for processing this manuscript.

My greatest thanks go to my dear wife Laurel and my wonderful daughters Amanda and Catherine for enduring, encouraging, and loving.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0 INTRODUCTION

Any transmission or reception at visible or infrared wavelengths may be adversely affected by cloudiness. For more than 20 years this has been the primary factor driving researchers toward accurate characterization of cloud-free regions.

Early investigators (McCabe, 1965; Lund, 1965; Lund and Shanklin, 1972, 1973; Rapp et al., 1973) used whole-sky photographs and/or cloud climatologies to estimate probabilities of cloud-free lines-of-sight from a specific point at the surface to a specific point in the atmosphere. No estimates of the extent of the cloud-free area surrounding the lines-of-sight were made. Grantham et al. (1979) addressed this deficiency. They detailed methods used to determine cloud-free fields-of-view. These methods were generalized to treat the case of an observer above the surface looking down as well as the previously treated case of the ground-based observer.

Interest in cloud-free regions as viewed from atop the atmosphere has grown as the importance of remote sensing by satellites has increased. But, satellite instruments do not record cloud-free regions in the same way as the human observer. Their radiometers scan one line at a time to produce an image of whatever lies below. Malick et al. (1979), imitating this scanning method, modeled probabilities of cloudy and clear intervals of specific lengths along straight-line paths. They accomplished this by thresholding satellite imagery to separate cloudy and clear regions and then constructing histograms of the cloudy and clear interval lengths. The resulting probability densities for both cloudy and clear lengths were distributed approximately exponentially with short intervals being much more likely than long intervals. Because the parameter used in the exponential probability density function is derived from the mean of the random variable in question, Malick et al. (1979) suggested use of the arithmetic means of cloudy and clear interval lengths to estimate those parameters.

Recent research has built upon the work of Malick et al. (1979). Goldstein and Janota (1985) provided a rigorous theoretical treatment of the exponential modeling of cloudy and clear intervals along finite paths. Snow et al. (1985b) and Snow (1985) used photographic imagery gathered during space shuttle orbiter missions to model interval length number densities with gamma, lognormal, and exponential functions. Snow and Willand (1986), again using space shuttle imagery, best fit interval length probabilities with the exponential probability density function. They also illustrated that cloud fraction values generated from very small portions of the original image exhibited little variablity, suggesting that image subsections could estimate the character of the entire image.

As mentioned, previous works were primarily concerned with the probability that a cloudy or clear interval equalled a specific length [Pr(X=x)]. A function used to describe such an equality is called a probability density function (PDF). A topic mentioned only in passing by Malick et al. (1979) was the probability that a cloudy or clear interval was *greater than* or equal to a specific length [Pr(X≥x)]. A function used to describe this is called, for want of a better term, one minus the cumulative distribution function (1-CDF) and is simply the integral of the PDF from x to infinity. These functions are simply alternate methods used to characterize cloudy and clear areas.

Characterizations of these types may be quite useful. An airborne or spaceborne observer might wish to view or transmit a signal to some point at the surface. If the observer needed a break in the clouds of at least one kilometer, he/she could calculate the clear interval 1-CDF and use it to predict a probability of success. The 1-CDF for clear intervals might also prove useful to navigational, earth resource, and cartographic satellite mission planners.

Atmospheric modelers might use such characterizations to improve parameterizations of cloudiness. Using a cloudy interval PDF would provide greater detail than the simple assignment of a particular fractional cloud cover. If the modeler were not interested in cloudiness, but rather in the distribution of moisture or vertical motion, cloudy intervals could be used as proxy variables to determine these distributions.

The purpose of this study, then, is to accurately model the PDFs and 1-CDFs of both cloudy and clear interval lengths within broken cloud fields. Three models are tested; their goodness-of-fit is evaluated with metric and congruent statistical methods not previously

used in these types of analyses. Image fractions as small as 1/32 are evaluated to determine how well they characterize the whole image.

The imagery used consists of 16 digitized cloud-field photographs taken during four space shuttle missions. This nearly doubles the number of cases considered to date in this line of research. This thesis constitutes the first formal report of its kind using very high resolution, photographic cloud imagery collected aboard the space shuttle.

2.0 DATA DESCRIPTION

2.1 DATA SOURCE

Space shuttle photographic images were chosen as data for this study because of their very high resolution and to illustrate their utility. Other data sources are available (e.g., Landsat) but are not commonly used to record cloud scenes.

The 16 digitized photographic images were collected during four space shuttle missions between September 1984 and October 1985. Both hand-held 35 millimeter and large format cameras were used. The hand-held camera photographs were collected as part of the Cloud Logic to Optimize Utilization of Defense Systems (CLOUDS) experiment. This experiment is sponsored by the United States Air Force and directed by the Air Force Geophysics Laboratory (AFGL). The large format camera photographs were collected as part of the Orbital Camera Payload System (OCPS) project. This project is co-sponsored by six United States Government agencies and directed by the Large Format Interagency Working Group. Magnetic tapes of these digitized images were provided by AFGL.

2.2 IMAGE SELECTION

The purpose of this study was to model cloudy and clear interval probability distributions in areas of broken cloudiness. For this reason, images with vast areas of overcast or completely clear skies were not used. The images were recorded from very close to the zenith to eliminate the apparent shortening of clear intervals and the distortion of cloudy intervals resulting from non-zero nadir angles. Althought many images were collected during the CLOUDS and OCPS projects, only 16 of them met both of these requirements. Fortunately, 11 of the 16 images chosen came from latitudes higher than any previously considered. The predominant cloud types in these images were cumulus and stratocumulus.

2.3 IMAGE COLLECTION AND CHARACTERISTICS

Five of the sixteen photographic images were taken with the hand-held camera on three separate shuttle missions. They were taken through the overhead flight deck windows. The remaining 11 photographic images were taken with the large format camera mounted in the shuttle cargo bay. Table 2.1 provides some particulars of the 16 photographs and Figure 2.1 shows the locations from which they were taken. Both the shuttle bay and overhead windows are directed earthward during normal shuttle operations.

A thorough treatment of image collection systems is provided by Snow and Tomlinson (1987). Information of primary interest to this study is shown in Table 2.2.

These image collection systems are non-standard (i.e., photographic) and record cloud scenes differently than satellite radiometers. Photographic emulsion charateristic curves relate the logarithm of film exposure to film density. Film exposure and scene reflectance are linearly related (Jensen, 1968) and film density is simply the logarithm of opacity. So the film characteristic curve, in the end, depicts the relationship between scene reflectance and film opacity. On the other hand, satellite radiometer calibration curves relate scene reflectance to reflectance count values.

Both of the films in Table 2.2 exhibit linear characteristic curves across the range of exposures used to collect imagery for this study. Film contrast is higher than in the original scene because the slopes of these curves exceed one. (The effect of this contrast enhancement on digitization is discussed in Section 3.1.)

TABLE 2.1   IMAGE SPECIFICATIONS

IMAGE NUMBER	LOCATION	LATITUDE (degrees north)	DATE/TIME** (YYMMDD/ HHHH)(GMT)	ORBITAL ALTITUDE (km)
1*	Dominican Republic	19	840902/1530	300
2	Bering Sea	57	841006/2100	270
3	Bering Sea	57	841006/2100	270
4	Bristol Bay	57	841006/2100	270
5	Bristol Bay	57	841006/2100	270
6	Bristol Bay	57	841006/2100	270
7	Bristol Bay	57	841006/2100	270
8	Bristol Bay	57	841006/2100	270
9	Bristol Bay	57	841006/2100	270
10	Poland	53	841007/1030	270
11	Poland	53	841007/1030	270
12	Poland	53	841007/1030	270
13*	Baja California Sur	25	850125/1700	300
14*	Florida	28	850125/1800	300
15*	Florida	27	851006/1400	450
16*	Florida	28	851006/1700	450

* These images were taken with the hand-held camera.
** All times are approximate.

(Of those images taken in the same place and at nearly the same time, no two are closer together than 75 kilometers.)

overhead windows

shuttle bay
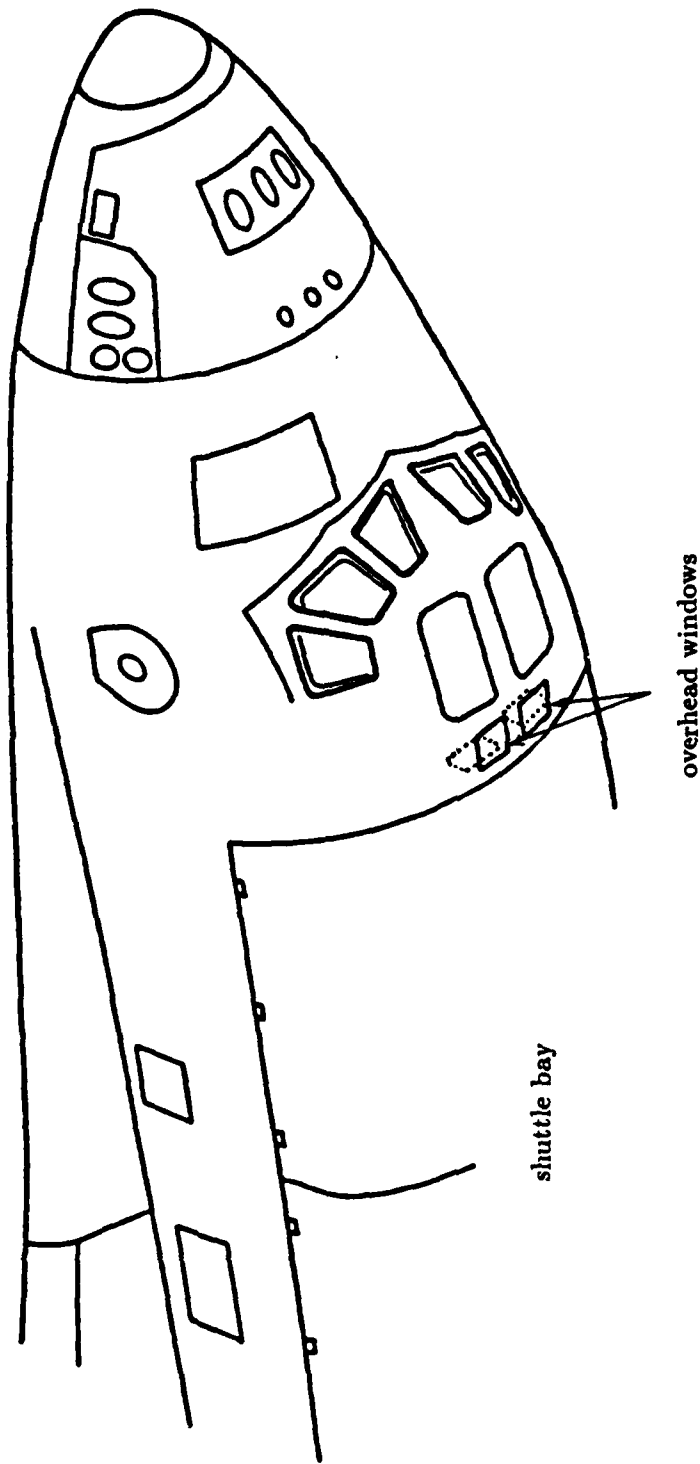
Figure 2.1   LOCATIONS FROM WHICH PHOTOGRAPHS WERE TAKEN

TABLE 2.2 IMAGE COLLECTION SYSTEMS CHARACTERISTICS

PROJECT NAME	CAMERA TYPE	LENS FOCAL LENGTH (mm)	LENS FIELD OF VIEW (degrees)	FILM TYPE	RESOLVING POWER* (lines/mm for high and low contrast)	IMAGE SIZE (mm x mm)	SPECTRAL RESPONSE AFTER FILTRATION** (μm)
CLOUDS	35mm	105	19.5	KODAK 2415	320 125	24 x 36	.60 to .70
OCPS	large format	305	41	KODAK 3414	125 90	229 x 457	.55 to .70

* Film resolving power is judged by the accuracy with which it reproduces the closely spaced parallel lines of a test chart.

** Filtration was used to reproduce the scatter of shorter wavelengths by atmospheric components.

3.0 DATA TREATMENT

3.1 IMAGE DIGITIZATION

To accomplish quantitative image analysis, each of the 16 images was digitized. Images 1, 13, and 14 were digitized by a private firm under contract with AFGL. The remaining 13 images were digitized at the Air Force Wright Aeronautics Laboratory (AFWAL). The original negative images were digitized using a linear relationship between negative opacity and digital count values. These count values were then expanded or contracted to match the 256 count range of the image display system. Reversal of the digital count values provided a positive digital image. Throughout the entire imaging process, linearity between the original scene and its digital representation was preserved although contrast was changed. This contrast manipulation did not affect brightness relationships within the image. An area appearing twice as bright as another in the original scene would still appear twice as bright in the final digitized image.

The extremely high resolution negative images were not digitized at full resolution due to equipment limitations and prohibitive expense. Table 3.1 lists the resolutions of the digitized images. Resolution of the digitized hand-held imagery was determined by counting the number of pixels constituting a surface feature of known dimension. In contrast, resolution of the digitized large format imagery was determined by knowing, in advance, the dimensions of the surface area covered by the imaging system. These dimensions were dictated by image collection system and orbital characteristics.

3.2 IMAGE CROPPING

The digitized images were displayed on the CIRA (C operative Institute for Research in the Atmosphere) COMTAL image processing system and reviewed for applicability to this study. The three images digitized under contract exhibited some noisy values near the

TABLE 3.1   DIGITIZED IMAGE SIZES AND RESOLUTIONS

IMAGE NUMBER	ORIGINAL IMAGE SIZE (pixels)	CROPPED IMAGE SIZE (pixels)	IMAGE RESOLUTION (m/pixel)	AREA OF CROPPED IMAGE (square km)
1	512 x 390	157 x 241	235	2089.548
2	2048 x 2048	512 x 345	99	1731.249
3	2048 x 2048	512 x 512	99	2569.273
4	2048 x 2048	512 x 512	49	629.408
5	2048 x 2048	512 x 512	49	629.408
6	2048 x 2048	310 x 240	49	178.634
7	2048 x 2048	512 x 512	49	629.408
8	2048 x 2048	512 x 512	49	629.408
9	2048 x 2048	512 x 512	99	2569.273
10	2048 x 2048	512 x 512	99	2569.273
11	2048 x 2048	512 x 512	99	2569.273
12	2048 x 2048	512 x 512	99	2569.273
13	270 x 256	241 x 240	300	5205.600
14	270 x 256	241 x 240	225	2928.150
15	2048 x 2048	260 x 365	65	400.953
15	2048 x 2048	463 x 373	58	580.959

image edges. The 13 images digitized by AFWAL, each 2048 x 2048 pixels, were too large to display on the image processing system. After eliminating noisy, extraneous, and large overcast and clear areas, 16 images of no greater size than 512 x 512 pixels remained. Table 3.1 provides the sizes of the original and cropped images.

3.3 IMAGE THRESHOLDING

After digitizing and cropping the images, choice of a threshold was necessary to distinguish cloudy and clear areas. Histograms of the first three images examined exhibited strong bimodality. The peak at lower reflectance values in each histogram corresponded to the background while the peak at higher reflectance values corresponded to areas well within clouds. A pronounced valley separated these two peaks. Table 3.2 shows at what reflectance values the two maxima and the dividing minimum occurred for images 1, 13, and 14. Figure 3.1 illustrates these features for image 13.

TABLE 3.2   HISTOGRAM MAXIMA AND MINIMA

IMAGE NUMBER	REFLECTANCE AT FIRST MAXIMUM	REFLECTANCE AT DIVIDING MINIMUM	REFLECTANCE AT SECOND MAXIMUM
1	74	128	186
13	9C	147	187
14	131	159	200

Knowing that the first peaks represented clear areas and the second peaks represented clouds, thresholding the images at the dividing minima was considered. In these cases, however, that tactic underestimated the actual cloud cover. The dividing minima values (and those just below them) corresponded to pixels within cloud edges. This is plausible when considering that cloud edges normally occupy less image area than either main cloud bodies or clear areas.
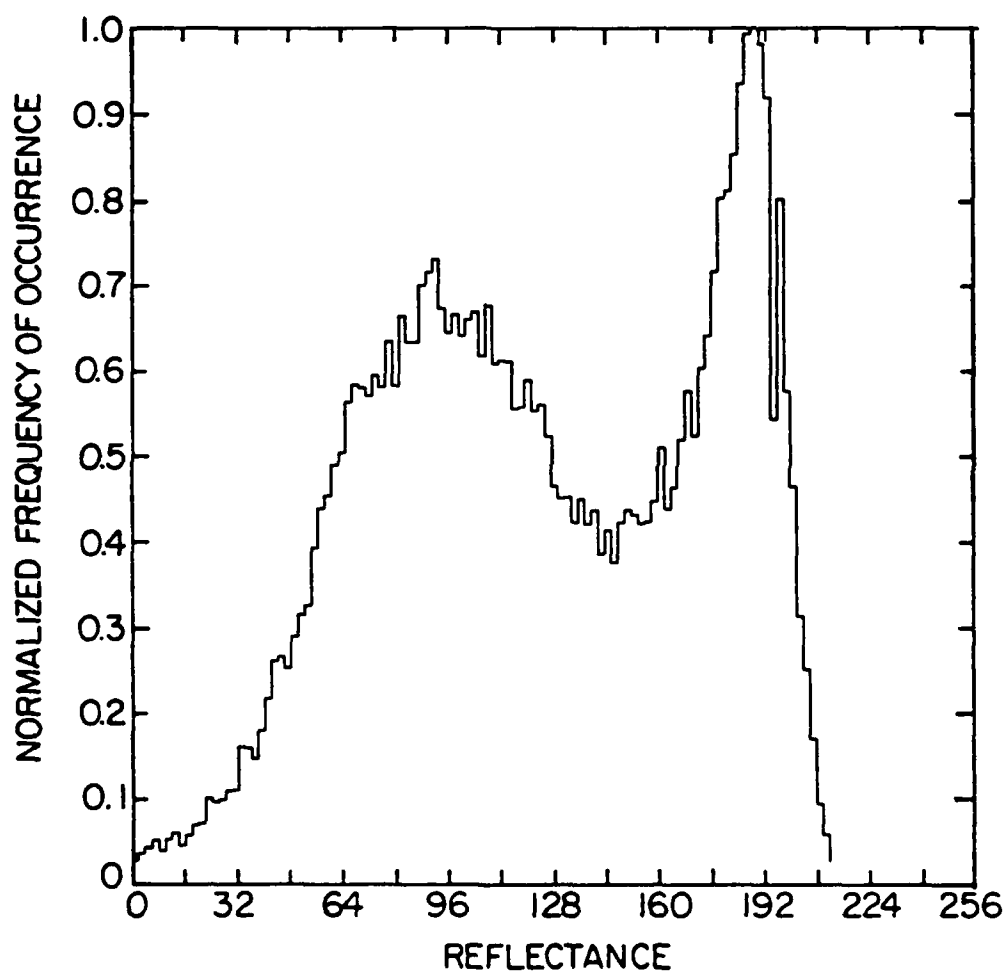
Figure 3.1   HISTOGRAM OF IMAGE 13

Having determined representative values for cloud edge and background reflectances, a thresholding technique discussed by Billingsley (1982) and Wielicki and Welch (1986) was chosen. The reflectance threshold, R_t, is given by

$$R_t = \frac{R_o + R_s}{2}, \tag{3.1}$$

where R_o is the average reflectance of a completely cloud filled pixel *within the cloud edge* and R_s is the clear-sky background reflectance.

This thresholding method worked very well for the three images listed in Table 3.2; however, most of the 13 remaining images did not exhibit bimodal reflectance histograms. In these, there were only pronounced maxima corresponding to the background reflectances, with higher count values nearly uniformly distributed. To solve the problem of locating the second peak in these histograms, smaller sections of each image that exhibited slightly higher cloudy fractions were examined. Figure 3.2 shows histograms of image 9 before and after inspection of a smaller cloudy section.

As shown in Figure 3.2b, these image sections did yield secondary histogram peaks corresponding to cloudy regions, but no sharp minima corresponding to cloud edges was present. This suggested determination of the Billingsley threshold in relation to the two histogram peaks rather than a first peak and dividing minimum.

Four sections of each bimodal image (i.e., a total of 12 subsections) were examined and the Billingsley threshold determined. The average fractional distance at which the thresholds occurred between the two maxima of each histogram was .254. The Billingsley threshold expression was then altered to read

$$R_t = R_s + .254(R_o - R_s). \tag{3.2}$$

Using this thresholding method, all 16 images were treated with consistency.

Histograms of no less than two image subsections were used to determine each image threshold. Subsection thresholds were averaged and applied to the whole. Pixels with reflectance at or above these thresholds were considered entirely cloudy and shaded white. Pixels with reflectance values below these thresholds were considered entirely clear and shaded black. All subsequent image manipulations were completed using these binary (i.e., black and white) images.

14
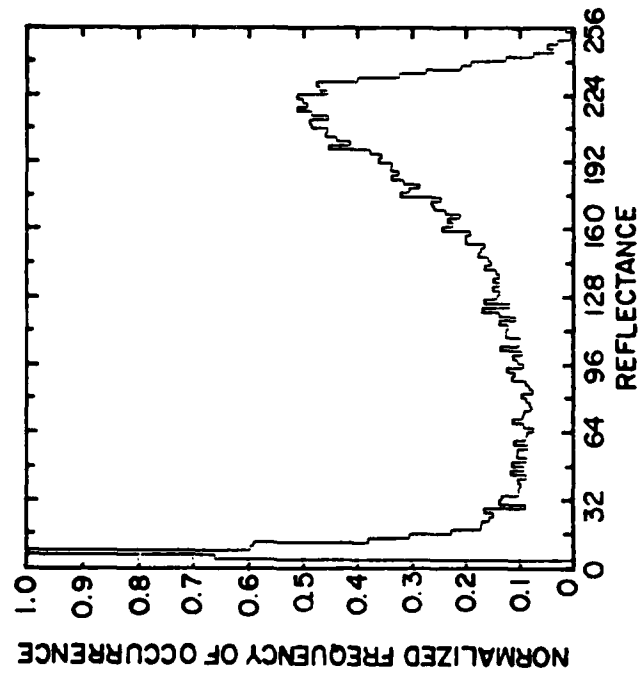
b) Section with slightly higher cloudy fraction
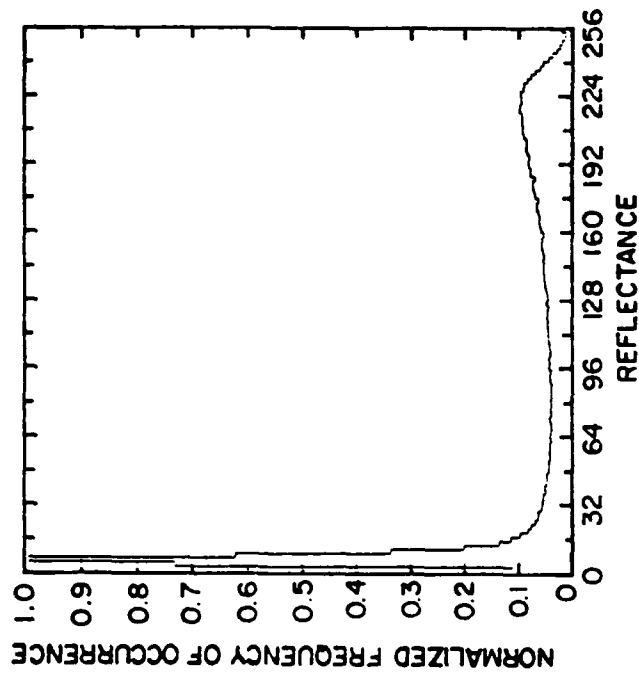
a) Whole image

Figure 3.2   HISTOGRAMS OF IMAGE 9

Appendix A contains copies of all 16 images, the subsections of those images treated in this study, and the image subsection after threshold application.

3.4 COUNTING ALGORITHM

The primary purpose of previous counting algorithms (Malick et al., 1979; Snow et al., 1985b) was to determine the probability that clear intervals of any given length existed in an image. (For simplicity, only the clear case is discussed at this point. Of course, any discussion applies equally to the cloudy interval lengths.) These algorithms were written to start at the upper left corner of an image and read from left to right until reaching the end of the line. The second line was then read from left to right and the program progressed in this fashion until the lower right corner of the image was reached. This treatment of the image was meant to simulate the scanning pattern used by satellite radiometers.

Counting all the way, the program tallied the number of clear intervals of each length encounterd. This information constituted the number density of the clear intervals. The number density values were then divided by the total number of clear intervals to yield the probability density. This represented $Pr(X=x)$ for clear interval lengths.

A problem occurs when intervals at the image edges are not carefully treated. These artificially shortened segments can bias any number or probability densities computed. In the described algorithms, these truncated intervals were tallieu along with all the others.

A primary aim in the design of this study's counting algorithm was to eliminate the inappropriate treatment of these artificially shortened intervals. This was accomplished by ignoring the first clear interval of each image line and any cloudy interval before it, and ignoring the last cloudy interval in each line and any clear interval following it. This ensured that the image used for statistical analysis consisted of an equal number of cloudy and clear intervals, none of which were artificially shortened. Having eliminated "undesirable" intervals the algorithm was used to generate the desired cloud field statistics.

To this point, with the exception of truncated interval treatment, this study's counting algorithm closely followed those used previously. But, in the present study the probability that a clear interval exceeded any given value $[Pr(X \geq x)]$ was also of interest. The program

used here just as easily tallied the numbers of clear intervals exceeding each possible value in the distribtion and then divided once more by the total number of clear intervals.

The primary output of the counting algorithm, then, was four files (two for the cloudy case and two for the clear case). The first and third files contained the probabilities that a randomly chosen cloudy or clear interval would equal any particular length [Pr(X=x)]. The second and fourth files contained the probabilities that a randomly chosen cloudy or clear interval would equal or exceed any chosen length [Pr(X≥x)]. Recall that functions used to model these probabilities are the PDF and 1-CDF.

The algorithm also computed the mean cloudy and clear interval lengths. To accomplish this, the total number of cloudy or clear pixels was divided by the number of cloudy or clear intervals.

3.5 MODEL CHOICES

Prospective models for the PDF and 1-CDF were chosen after viewing plots (on linearly scaled paper) of the actual probability distributions. Figure 3.3 shows these plots for cloudy intervals of image 1.

Plots for the other 15 images were quite similar in shape with short interval length probabilities considerably higher than those of long intervals. Contributions to these high probabilities were made by both small clouds and small appendages extending from larger clouds.

After viewing all the plots, model choices were narrowed to exponential, hyperbolic, and logarithmic functions. These functions were particularly convenient to work with because they are easily linearized, simplifying subsequent regressions. Table 3.3 shows the functions and their linear transforms. In Table 3.3, x represents the chosen cloudy interval length and Y represents the modeled Pr(X=x) or Pr(X≥x), depending on whether the PDF or 1-CDF is being considered.

Only β needed to be estimated before model probabilities were compared with the observed values. To estimate β, actual values of Y and X (or functions of these) were compared by linear regression using the linear transforms of the 1-CDFs shown in Table 3.3. The regression algorithm minimized the sum of the *absolute* deviations about the regression

a) Probability that a randomly chosen
   cloudy interval equals x

b) Probability that a randomly chosen
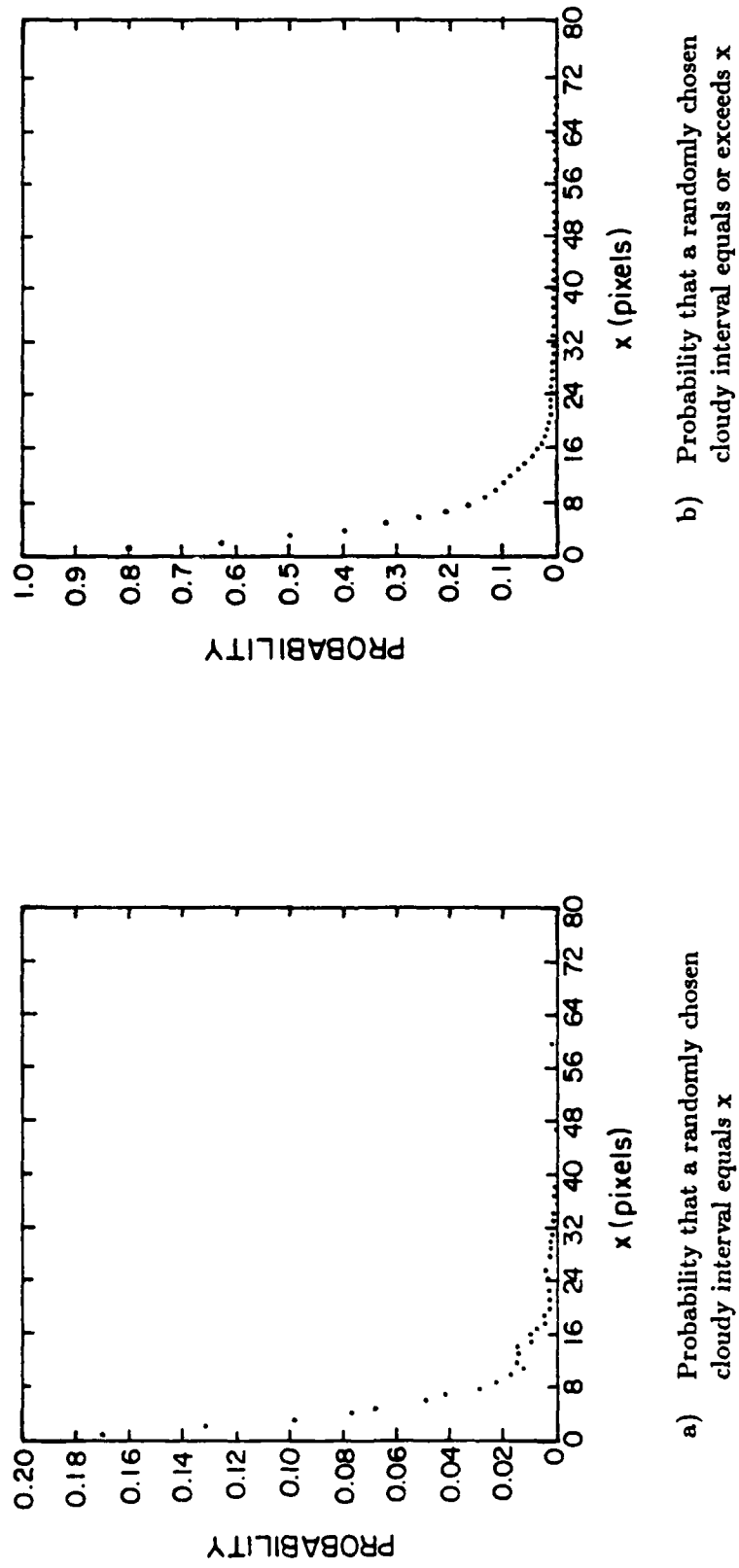   cloudy interval equals or exceeds x

Figure 3.3   PROBABILITY DISTRIBUTIONS FOR IMAGE 1

TABLE 3.3   PROSPECTIVE MODEL FUNCTIONS

MODEL*	PDF	PDF (linear transform)	1-CDF	1-CDF (linear transform)
exp	$Y=-\beta\exp(\beta x)$	$\ln(Y)=\ln(-\beta)+\beta x$	$Y=\exp(\beta x)$	$\ln(Y)=0+\beta x$
hyp	$Y=-\beta\exp(\beta-1)$	$\ln(Y)=\ln(-\beta)+(\beta-1)\ln(x)$	$Y=x^\beta$	$\ln(Y)=0+\beta\ln(x)$
log	$Y=-\beta/x$	$Y=0+(-\beta/x)$	$Y=\beta\ln(x)$	$Y=0+\beta\ln(x)$

*   exp = exponential, hyp = hyperbolic, log = logarithmic

line to determine β rather than minimizing the sum of the *squared* deviations. (NOTE: Because the βs were determined using transforms of the original equations, there was no guarantee that errors would still be minimal when respective β values were used in the original equations.)

3.6   GOODNESS-OF-FIT

Values of β determined by regression were then placed into the original model equations for PDF and 1-CDF and compared to the observed probabilities. This comparison was made using a goodness-of-fit statistic based on least absolute deviation (See Appendix B.). The goodness-of-fit statistic, G, was computed by

$$G = (1/n)\sum_{i=1}^{n} |Y - \hat{Y}|, \tag{3.3}$$

where n is the number of observations (i.e., length categories), Y is the observed probability value, and $\hat{Y}$ is the model estimate of the observed value.

Table 3.4 shows the number of observations used for each model-to-observation comparison. As an illustrative example, consider image 1. Cloudy intervals range in length from 1 to 71 pixels and clear intervals from 1 to 49 pixels. These are the number of cases for which values of the 1-CDF can be computed. But when the PDF is considered for each case, this number decreases. Even though there are cloudy intervals from as short as 1

TABLE 3.4 NUMBER OF OBSERVATIONS USED IN COMPARISONS

IMAGE NUMBER	Pr(X=x) OR PDF FOR CLOUDY CASE	Pr(X≥x) OR 1-CDF FOR CLOUDY CASE	Pr(X=X) OR PDF FOR CLEAR CASE	Pr(X≥X) OR 1-CDF FOR CLEAR CASE
1	38	71	38	49
2	110	137	133	218
3	72	104	241	418
4	134	173	167	413
5	163	269	121	141
6	54	77	63	114
7	161	215	181	287
8	127	216	200	275
9	159	193	112	159
10	91	132	223	418
11	94	115	229	352
12	83	90	150	178
13	45	50	42	58
14	104	179	41	89
15	91	116	76	105
16	48	66	134	254
TOTAL*	1574	2203	2151	3528

* "n" is equal to these values in computation of the
G statistic for the 16 images as a group. Other
values of "n" shown here are used in Appendix C
to compute G values for individual images.

pixel to as long as 71 pixels, not every length category between these is represented by a cloudy length. In the case of image 1 there are 33 cloudy lengths (71 minus 38) that are not represented. For the clear case there are only 11 lengths (49 minus 38) missing. These are simply the unobserved lengths between shortest and longest observed values. In image 1 there are no clear intervals 29, 32, 35, 38, 40, 41, 44, 45, 46, 47, or 48 pixels long. Errors for these non-existent lengths were not computed when determining the goodness-of-fit of the PDF models.

4.0 RESULTS

4.1 MODEL PERFORMANCE

Values of G for model treatments of individual images are presented in Appendix C while those for the 16 images as a group are presented in Table 4.1.

The exponential model exhibited a smaller mean error than the other two models in every condition/case category. This agreed with the findings of Snow and Willand (1986), who found that the exponential distribution outperformed the lognormal distribution and two varieties of the gamma distribution. (Their results were based only on fits of PDFs, however.) Further examination of the exponential model was merited on the basis of these small errors.

4.2 EXPONENTIAL MODEL INVESTIGATION

Malick et al. (1979) were the first to suggest that the mean interval length (cloudy or clear) be used to estimate β when using the exponential distribution functions. Snow and Willand (1986) were the first to use this concept on actual data. They estimated β for the cloudy case by multiplying the cloud cover fraction by the recurrence interval (simply the sum of the mean cloudy and clear interval lengths), realizing that the result was identical to the reciprocal of the mean cloudy interval length. For the clear case, they used the clear fraction times the recurrence length.

The counting algorithm used in this study computed both the mean cloudy and mean clear interval lengths for each image. The reciprocals of these are the actual estimators of β. The resulting values were substituted into the original exponential equations to determine whether or not mean errors could be reduced.

Table 4.2 is a comparison of exponential model results using βs estimated by both regression and cloud-field statistics (mean cloudy and clear interval lengths).

TABLE 4.1 MEAN ERRORS OF MODELED PROBABILITIES

CONDITION	CASE	MODEL	TOTAL ERROR	NUMBER OF OBSERVATIONS	G^* (mean error)
CLOUDY	Pr(X=x) OR PDF	exp	6.1494	1574	.00391**
		hyp	24.9089	1574	.01583
		log	16.3421	1574	.01038
	Pr(X≥x) OR 1-CDF	exp	74.0809	2203	.03363**
		hyp	145.0984	2203	.06584
		log	223.3028	2203	.10136
CLEAR	Pr(X=x) OR PDF	exp	8.6670	2151	.00403**
		hyp	23.2122	2151	.01079
		log	16.3199	2151	.00759
	Pr(X≥x) OR 1-CDF	exp	165.3093	3528	.04686**
		hyp	209.6766	3528	.05943
		log	313.2807	3528	.08880

* Multiply these values by 100 to find the percent error.

** These models performed best in their respective CONDITION/CASE combinations.

TABLE 4.2 ERRORS USING β FROM REGRESSION AND FROM CLOUD FIELD

CONDITION	CASE	METHOD*	TOTAL ERROR	NUMBER OF OBSERVATIONS	G^{**} (mean error)	ERROR*** REDUCTION (percent)
CLOUDY	Pr(X=x)	regre	6.1494	1574	.00391	
	OR					23
	PDF	cloud	4.7423	1574	.00301	
	Pr(X≥x)	regre	74.0809	2203	.03363	
	OR					43
	1-CDF	cloud	42.1936	2203	.01915	
CLEAR	Pr(X=x)	regre	8.6670	2151	.00403	
	OR					20
	PDF	cloud	6.9140	2151	.00321	
	Pr(X≥x)	regre	165.3093	3528	.04686	
	OR					39
	1-CDF	cloud	100.3047	3528	.02843	

* regre = determined from regression
cloud = determined from cloud-field statistics
** Multiply these values by 100 to find the percent error.
*** All error reduction resulted from use of β determined using
cloud-field statistics (mean cloudy and clear interval lengths).

It is at first surprising that cloud-field statistical estimates of β perform better than those estimated by regression. The apparent contradiction is rectified by recalling that the regression minimized error for the transformed equations, not the original equations. (See NOTE at the end of Section 3.5.)

4.3 IMAGE FRACTION STATISTICS

How small can an image fraction be while still accurately reflecting the probability distributions of the original image? Snow and Willand (1986) provided part of the answer. They treated images by looking at short strings of pixels, one after another, and determining the variance of both cloudy fraction and recurrence length. They systematically increased the pixel string lengths until the variance stabilized at a low value. For the few images they considered, the variance reached and maintained a minimum when the image was sampled in successive strings of about 250 or more pixels. This showed there was little variance in cloud cover and recurrence length from one 250 pixel string to the next. This variability, even though low, was still based on examination of the entire image and did not consider the result of examining only one (perhaps the first) 250 pixel string.

This study considered smaller and smaller fractions of each image in preference to pixel strings of fixed length. This was done to determine just what fraction of an image, rather than absolute number of pixels, needs to be sampled to accurately represent the entire image in terms of its mean cloudy or clear interval length. Initially, the first half of an image was considered; then, the first quarter. This halving of each successive fractional portion continued until the first 1/32 of each image was treated. Each of these fractions was taken from the beginning of the image (as a satellite might view it when passing overhead) so that the resulting model errors would reflect the inaccuracies of a spatial forecast. Values of β were determined for each fractional image and then used in the exponential model to estimate the probability distributions of the original, whole image. Mean errors resulting from use of the fractional β values were then computed.

Table 4.3 presents the mean errors for fractional images. As expected, errors increased as smaller and smaller image fractions were used to model the entire image. The subsequent halving of the image fractions was stopped when two successive image fractions produced

TABLE 4.3   MEAN ERRORS ASSOCIATED WITH FRACTIONAL IMAGES

CONDITION	CASE	IMAGE FRACTION CONSIDERED	G^* (mean error)
CLOUDY	Pr(X=x)	1/1	.00301
	OR	1/2	.00325
	PDF	1/4	.00345
		1/8	.00371
		1/16	.00375
		1/32	.00400
	Pr(X≥x)	1/1	.01915
	OR	1/2	.02492
	1-CDF	1/4	.03576
		1/8	.04993
		1/16	.05678
		1/32	.06274
CLEAR	Pr(X=x)	1/1	.00321
	OR	1/2	.00389
	PDF	1/4	.00358
		1/8	.00336
		1/16	.00359
		1/32	.00408
	Pr(X≥x)	1/1	.02843
	OR	1/2	.03668
	1-CDF	1/4	.05223
		1/8	.03684
		1/16	.05128
		1/32	.05914

* Multiply these values by 100 to find the percent error.

mean errors exceeding five percent. This occurred at an image fraction of 1/32 for both cloudy and clear 1-CDFs. Errors for the PDFs were considerably smaller than the 1-CDFs because the estimated probabilities were, themselves, smaller.

5.0 SUMMARY AND CONCLUSIONS

5.1 SUMMARY

This study used 16 very high resolution, digitzed photographic images collected during 4 space shuttle orbiter missions. They were used to determine a realistic model of the PDF and 1-CDF of cloudy and clear interval lengths within broken cloud-fields. Three models were tested: exponential, hyperbolic, and logarithmic.

Image fractions as small as 1/32 were used to estimate the character of the entire cloud-field. These spatial forecasts and the interval probability models were tested using statistical methods based on least absolute deviations.

5.2 CONCLUSIONS

The major conclusions of this study are as follows:

1. The use of the dividing minimum reflectance value, between the two bimodal histogram maxima of a visible image, as a threshold between cloudy and clear regions is questionable. Pixels of this and slightly lower reflectance values fell within cloud edges causing many cloudy pixels to be classified as clear. A thresholding technique loosely based on that of Billingsley (1982) was proposed. It placed the threshold at approximately one quarter of the distance between the low and high value peaks of the histogram. (The technique worked very well in this study; however, it was developed using only a small number of very high resolution visible images.)

2. Exponential models can accurately represent observed probabilities of cloudy and clear interval lengths within cumuliform cloud fields (see Table 4.1). Very good estimates of the exponential distribution parameters are provided simply by the inverses of the mean cloudy and mean clear interval lengths (see Table 4.2).

3. Image fractions as small as the first 1/32 can estimate, with reasonable accuracy, the PDF and 1-CDF of the cloudy and clear interval lengths of the entire image (see Table 4.3). In other words, a form of spatial forecast can be made of what probabilities will occur in an area larger than that actually sampled.

Conclusions two and three agree with and extend the findings of Snow and Willand (1986). It is encouraging, in light of dissimilar statistical and image edge treatments, that this and previous works agree on the utility of the exponential model in representation of cloudy and clear interval length probabilities. However, the underlying physical principle dictating the distribution of these interval lengths remains unexplained.

If interval lengths are considered a measure of the entropy or degree of order of the cloud system, an argument made by Levine (1983) may be used. He proposed that entropy be defined as a function of the probability of a state. He proved that this function of probability must take the form

$$S = k[ln(Pr)] + a, \tag{5.1}$$

where S is entropy, k and a are constants, and Pr is the probability of the state. It is quite interesting to find that substitution of Y and x (from section 3.5) for Pr and S of (5.1) yields an equation of the same form as the exponential PDF and 1-CDF.

5.3 POSSIBLE APPLICATIONS AND FUTURE WORK

5.3.1 Applications

Although many applications of this study's results may exist, those in the atmospheric sciences are of primary interest.

Atmospheric boundary layer characteristics control the growth of cumuliform clouds. From one location to another and from one time to another these characteristics change, with no two boundary layer air masses entirely identical. If clouds are present, the mean cloudy and clear interval lengths also change. It may be possible to equate different values of β (or its reciprocal) with specific air mass types or with single air masses during different phases of boundary layer development. This would constitute a simple parameterization of boundary layer characteristics that might prove useful in atmospheric models of sufficiently

high resolution. The cloudy and clear interval lengths could be treated as proxies for the atmospheric variables associated with them. This could allow a modeler to dictate the distribution of moisture, vertical motion, etc. within the model.

Modelers of the atmospheric boundary layer and convective processes might use the exponential PDF as a necessary but insufficient condition for cloud-field realism. Cloud scenes generated by their models could be subjected to the counting algorithm proposed in this paper. Resulting mean cloudy and clear lengths could be used to generate the PDFs that the model cloud-field *should* exhibit. The more closely the model probabilities match the exponential PDFs, the better the model representation of reality (i.e., reality as the exponential PDFs represents it). While this constitutes a comparison of two models, one theoretical and one empirical, the exponential model has thus far proven quite accurate in its representations of cloud scenes.

5.3.2 Future Work

The space shuttle orbiter is capable of collecting imagery at extremely high resolutions (see Appendix D). These data could be used to examine atmospheric phenomena with a combination of global coverage and detail not previously possible. (Satellites may have greater global coverage and aerial photographs may have greater spatial resolution, but the space shuttle provides an unparallelled combination of the two.)

Cumulus and stratocumulus were the predominant cloud types in this study. Other broken cloud types should be examined to determine whether or not their interval length probability distributions can be described by exponential functions. It may also be useful to describe how these distributions change as azimuthal scanning angle is varied for single images.

Currently, accurate description of cloud size distributions is possible, but two-dimensional representation of spaces separating those clouds is difficult. Consequently, anyone interested in cloud-field representations may know the sizes of clouds but not know how far apart to place them. This study has shown that both cloudy and clear spaces can be accurately characterized in one dimension. If one- and two-dimensional representations can

be directly related, perhaps more accurate two-dimensional cloud-field chararcterizations could be accomplished.

REFERENCES

Agee, E. M., 1984: Observations from Space and Thermal Convection: A Historical Perspective. *Bull. Amer. Meteor. Soc.*, 65, 938–949.

Allen, J. H., and J. D. Malick, 1983: The Frequency of Cloud-Free Viewing Intervals. American Institute of Aeronautics and Astronautics, Inc., 5pp.

Barrodale, I., and F. D. K. Roberts, 1973: An Improved Algorithm for Discrete L1 Linear Approximation. *SIAM J. Numer. Anal.*, 10, 839–848.

Behunek, J. L., T. H. Vonder Haar, and M. A. Klitch, 1982: Satellite-Derived Cloud Statistics for Great Plains Cumulus. Presentations at the Conference on Cloud Physics. American Meteorological Society, 263–266.

Billingsley, F. C., 1982: Modeling Misregistration and Related Effects on Multispectral Classification. *Phot. Eng. Remote Sens.*, 48, 421–430.

Blackmer, R. H., Jr., and S. M. Serebreny, 1962: Dimensions and Distributions of Cumulus Clouds as Shown by U-2 Photographs. Stanford Research Institute, AFCRL-62-609, 57pp.

Bloomfield, P., and W. L. Steiger, 1983: Least Absolute Deviations. Birkhauser Boston, Inc., 349pp.

Claerbout, J. F., and F. Muir, 1973: Robust Modeling with Erratic Data. *Geophys.*, 5, 826–844.

Coakley, J. A., Jr., and F. P. Bretherton, 1982: Cloud Cover from High-Resolution Scanner Data: Detecting and Allowing for Partially Filled Fields of View. *J. Geophys. Res.*, 87, 4917–4932.

Cox, A., 1974: Photographic Optics. Focal Press, 596 pp.

Devore, J. L., 1982: Probability and Statistics for Engineering and the Sciences. Brooks/Cole Publishing Company, 640pp.

Feigelson, E. M., 1984: Radiation in a Cloudy Atmosphere. D. Reidel Publishing Company, 293 pp.

Frank, W. M., 1983: The Cumulus Parameterization Problem. *Mon. Wea. Rev.*, 111, 1859–1871.

Goldstein, J. D., and P. Janota, 1985: A Probabilistic Model for the Length of Cloud-Free Intervals Along Finite Paths. Presentations at the Third Tri-Service Cloud Modeling Workshop. Air Force Geophysics Laboratory, 219–226.

Grantham, D. D., I. A. Lund, and R. E. Davis, 1979: Estimating the Probability of Cloud-Free Fields-of-View Between Earth and Airborne or Space Platforms. USAF Air Weather Service Technichal Report 79/001, 207–214.

Greaves, J. R., P. E. Sherr, and A. H. Glaser, 1970: Cloud Cover Statistics and Their Use in the Planning of Remote Sensing Missions. *Remote Sens. Environ.*, 1, 95–101.

Hozumi, K., T. Harimaya, and C. Magono, 1982: The Size Distribution of Cumulus Clouds as a Function of Cloud Amount. *J. Meteor. Soc. Japan*, 60, 691–699.

James, T. H., and G. C. Higgins, 1948: Fundamentals of Photographic Theory. John Wiley and Sons, Inc., 286 pp.

Jensen, N., 1968: Optical and Photographic Reconnaissance Systems. John Wiley and Sons, Inc., 211 pp.

Kauth, R. J., and J. L. Penquite, 1967: The Probability of Clear Lines of Sight Through a Cloudy Atmosphere. *J. Appl. Meteor.*, 6, 1005–1017.

Levine, I. N., 1983: Physical Chemistry. McGraw-Hill, Inc., 890 pp.

Lopez, R. E., 1977: The Lognormal Distribution and Cumulus Cloud Populations. *Mon. Wea. Rev.*, 105, 865–872.

Lund, I. A., 1965: Estimating the Probability of Clear Lines-of-Sight from Sunshine and Cloud Cover Observations. *J. Appl. Meteor.*, 4, 714–722.

Lund, I. A., and M. D. Shanklin, 1972: Photogrammetrically Determined Cloud-Free Lines-of-Sight Through the Atmosphere. *J. Appl. Meteor.*, 11, 773–782.

Lund, I. A., and M. D. Shanklin, 1973: Universal Methods for Estimating Probabilities of Cloud-Free Lines-of-Sight Through the Atmosphere. *J. Appl. Meteor.*, 12, 28–35.

Lund, I. A., D. D. Grantham, and R. E. Davis, 1980: Estimating Probabilities of Cloud-Free Fields-of-View from the Earth Through the Atmosphere. *J. Appl. Meteor.*, 19, 452–463.

Malick, J. D., J. H. Allen, and S. Zakanycz, 1979: Calibrated Analysis of Cloud-Free Intervals. Proceedings of SPIE, 195, 142–147.

McCabe, J. T., 1965: Estimating Mean Cloud and Climatological Probability of Cloud-Free Line-of-Sight. USAF Air Weather Service Technical Report 186, 26pp.

McAllister, C. R., 1969: Cloud-Cover Recurrence and Diurnal Variation. *J. Appl. Meteor.*, 8, 769–775.

Mielke, P. W., Jr., 1986: Non-Metric Statistical Analyses: Some Metric Alternatives. *J. Statistical Planning and Inference*, 13, 377–387.

Narula, S. C., and J. F. Wellington, 1982: The Minimum Sum of Absolute Errors Regression: A State of the Art Survey. *Int. St. Rvw.*, 50, 317–326.

Parker, L., R. M. Welch, and D. J. Musil, 1986: Analysis of Spatial Inhomogeneities in Cumulus Clouds Using High Spatial Resolution Landsat Data. *J. Climate Appl. Meteor.*, 25, 1301–1314.

Plank, V. G., 1969: The Size Distribution of Cumulus Clouds in Representative Florida Populations. *J. Appl. Meteor.*, 8, 46–67.

Rapp, R. R., C. Schultz, and E. Rodriguez, 1973: Cloud-Free Line-of-Sight Calculations. *J. Appl. Meteor.*, 12, 484–493.

Raymond, K. W., N. Chidambaram, and P. W. Mielke, Jr., 1983: Application of Multi-Response Permutation Procedures and Median Regression for Covariate Analyses of Possible Weather Modification Effects on Hail Responses. *Atmosphere-Ocean*, 21, 1–13.

Shenk, W. E., and V. V. Salomonson, 1972: A Simulation Study Exploring the Effects of Sensor Spatial Resolution on Estimates of Cloud Cover from Satellites. *J. Appl. Meteor.*, 11, 214–220.

Snow, J. W., 1985: Some Cloud Population Statistics. Proceedings NASA Symposium on Global Wind Measurements. Columbia, MD., 229–234.

Snow, J. W., J. T. Bunting, R. P. d'Entremont, D. D. Grantham, K. R. Hardy, E. M. Tomlinson, 1985a: Space Shuttle Cloud Photographs Assist in Correcting Meteorological Satellite Data. *EOS*, 66, 489–490.

Snow, J. W., E. Tomlinson, and J. H. Willand, 1985b: Distributions of Clear and Cloudy Intervals, Spacial Correlations from Space Shuttle Photographs. Presentations at the Third Tri-Service Cloud Modeling Workshop. Air Force Geophysics Laboratory, 75–89.

Snow, J. W., and J. H. Willand, 1986: Computing Clear-Interval Probabilities for Cloud Scenes Photographed from the Space Shuttle. Presentations at the Fourth Tri-Service Cloud Modeling Workshop. Air Force Geophysics Laboratory, 184–198.

Snow, J. W., and E. M. Tomlinson, 1987: Cloud Population Measurements Using Photographs Taken from the Space Shuttle. Proceedings 6th Symposium on Meteorological Observations and Instrumentation. AMS, New Orleans, LA., 286–289.

Walls, L. W., 1973: The Beta Distribution: A Statistical Model for World Cloud Cover. NASA Technical Memorandum X-64714, 53pp.

Wielicki, B. A., and R. M. Welch, 1986: Cumulus Cloud Properties Derived Using Landsat Satellite Data. *J. Climate Appl. Meteor.*, **25**, 261–276.

APPENDIX A

IMAGERY:

This section provides the reader with prints of the imagery used in this study. The prints were made using a Talaris 800 Laserprinter. The input to the printer was provided by the CIRA COMTAL image processing system. Having undergone several manipulations, the images no longer display their orignal resolution. The resolution of the printed images is four times more coarse than when they were displayed on the COMTAL imaging system.

Each print is divided into four quadrants. The contents of each quadrant is explained below.

UPPER LEFT QUADRANT: This image is shown as it was displayed on the COMTAL imaging system. For the 2048 x 2048 pixel images, this is the 512 x 512 section selected from it.

UPPER RIGHT QUADRANT: This image is the cropped image and is a subsection of the image in the upper left quadrant. If the image was not cropped, it appears exactly as it did in the upper left quadrant.

LOWER RIGHT QUADRANT: This image exactly matches the one above it except that the cloudy/clear threshold has been applied. Clouds appear white and clear areas appear black.

LOWER LEFT QUADRANT: Along with pixel resolution, here are some image statistics not presented elsewhere in the paper.

resolution:
235 meters/pixel

cropped image size:
36.9 x 56.6 km

mean cloudy length:
1.23 km

mean clear length:
1.08 km
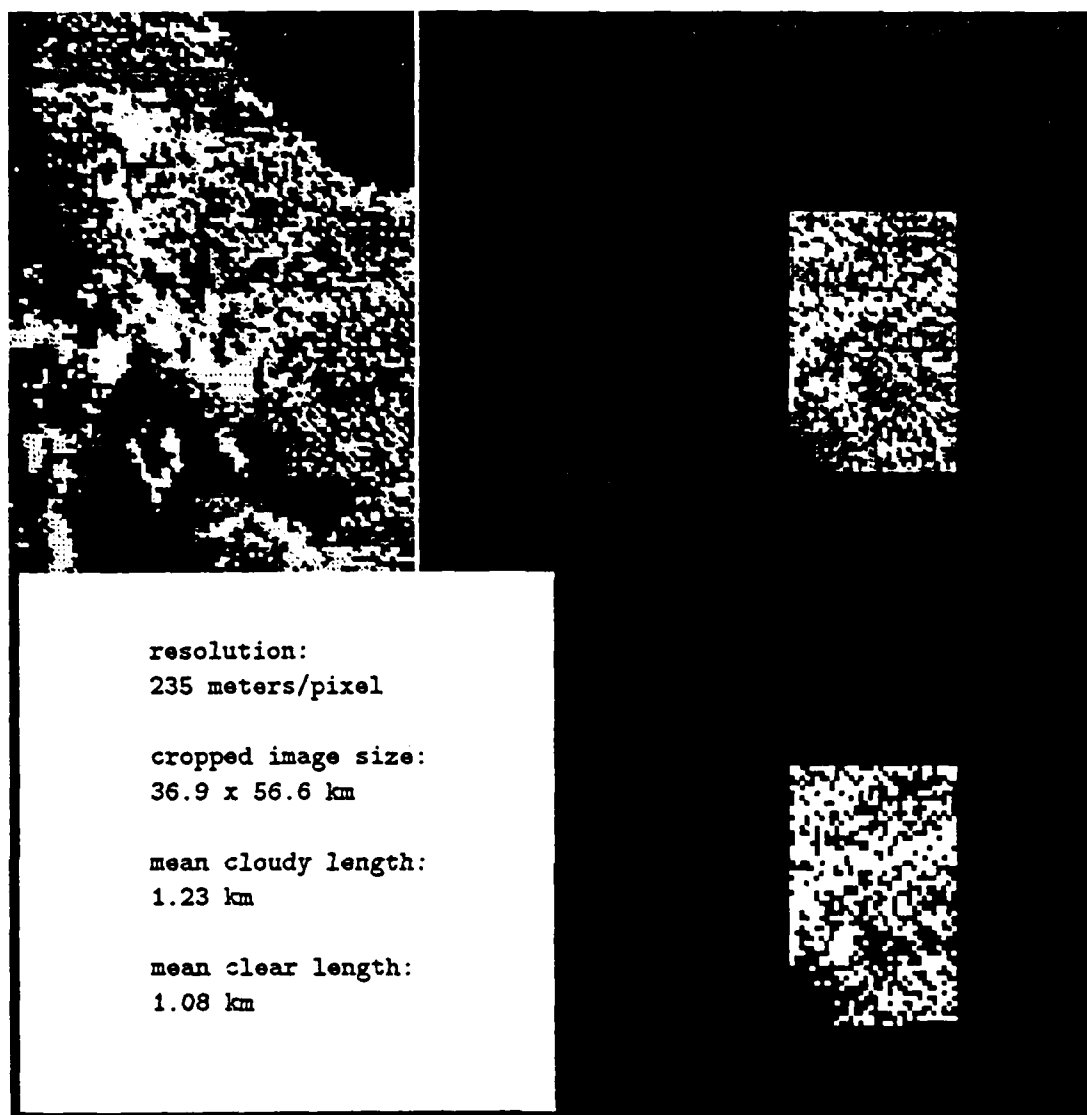
Figure A.1   TREATMENTS AND STATISTICS OF IMAGE 1
(See Appendix A text for figure explanation.)

resolution:
99 meters/pixel

cropped image size:
50.7 x 34.2 km

mean cloudy length:
1.85 km

mean clear length:
2.50 km
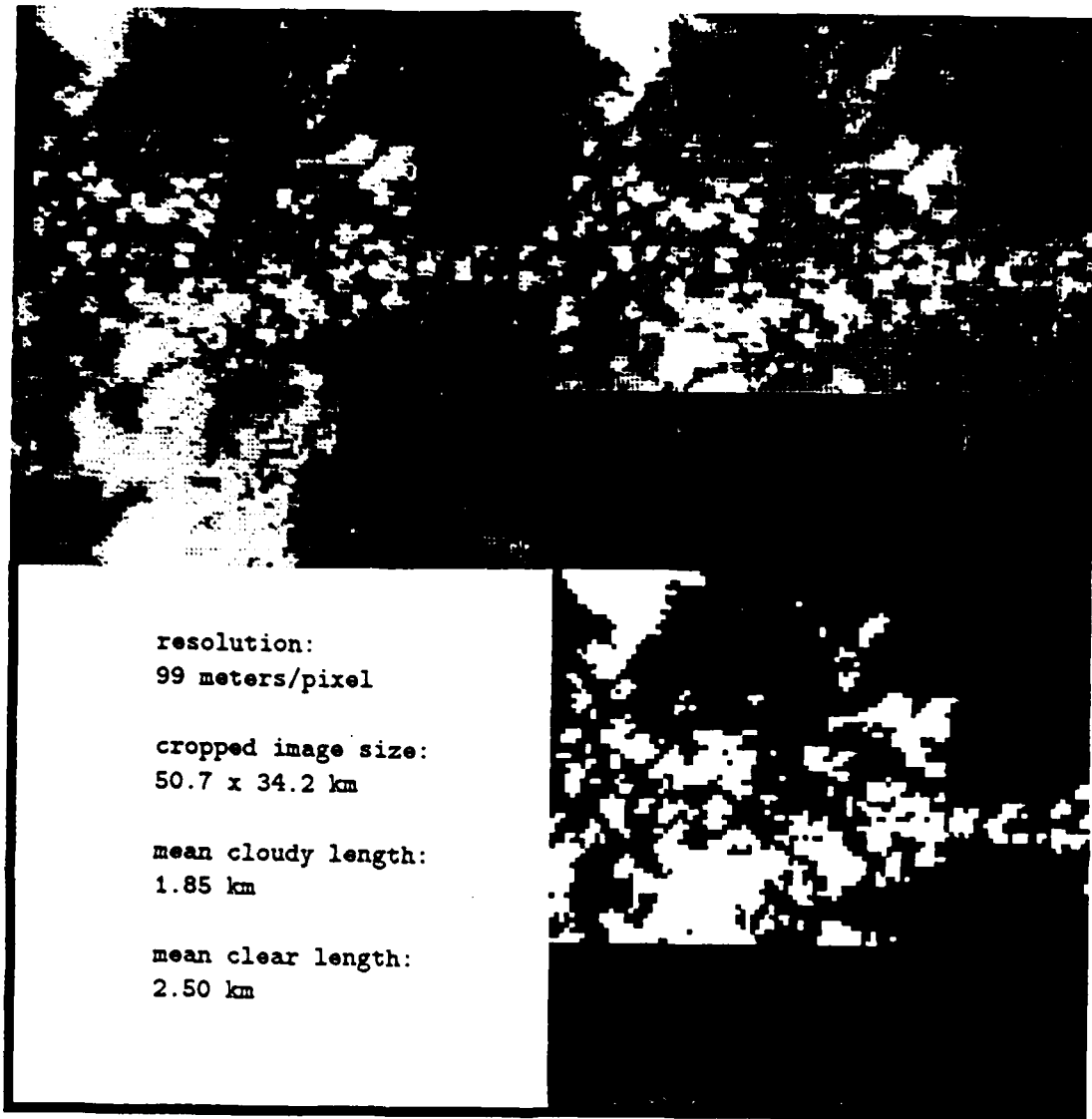
Figure A.2   TREATMENTS AND STATISTICS OF IMAGE 2
(See Appendix A text for figure explanation.)

resolution:
99 meters/pixel

cropped image size:
50.7 x 50.7 km

mean cloudy length:
1.29 km

mean clear length:
5.01 km
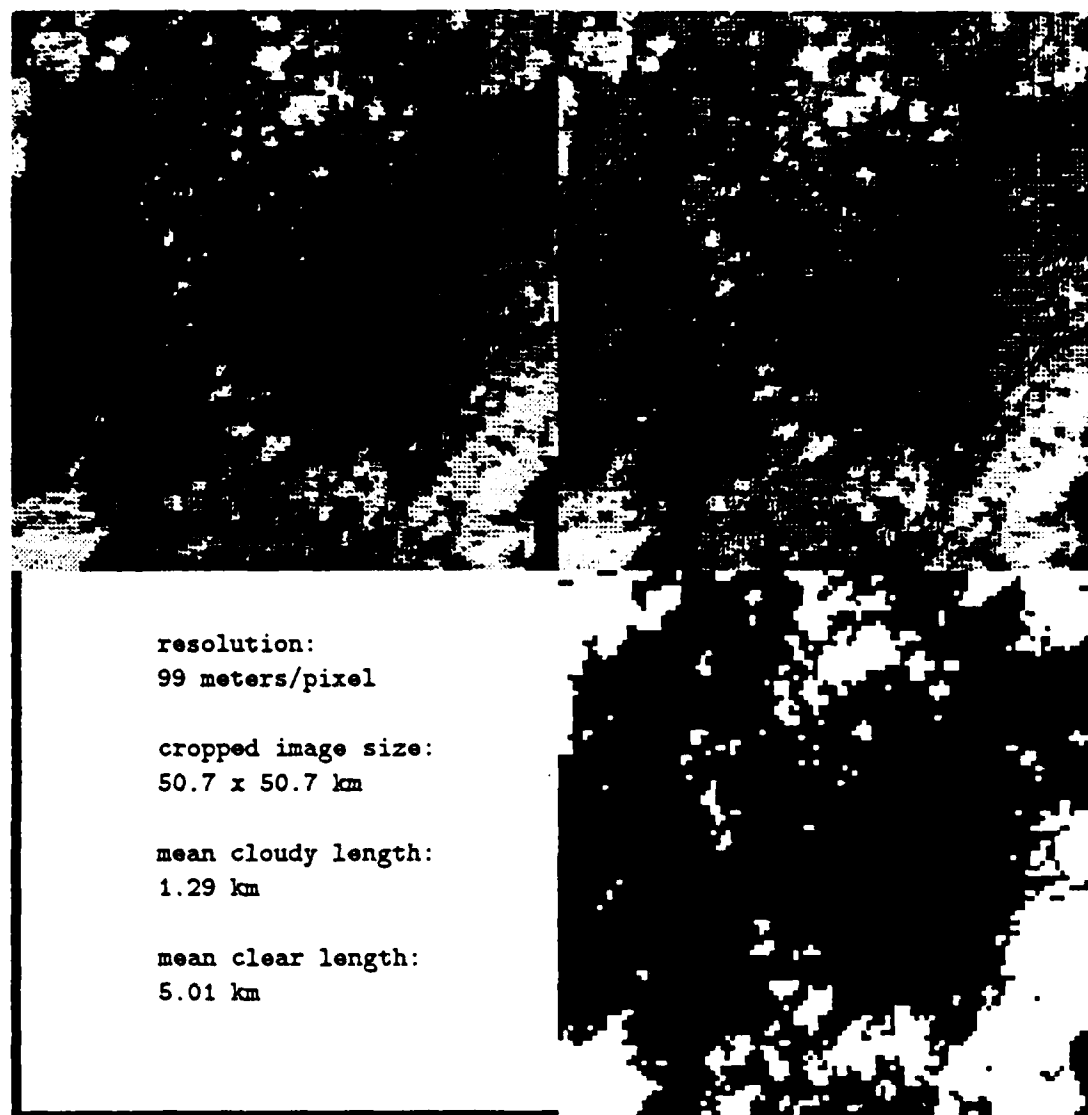
Figure A.3   TREATMENTS AND STATISTICS OF IMAGE 3
(See Appendix A text for figure explanation.)

resolution:
49 meters/pixel

cropped image size:
25.1 x 25.1 km

mean cloudy length:
0.87 km

mean clear length:
1.05 km
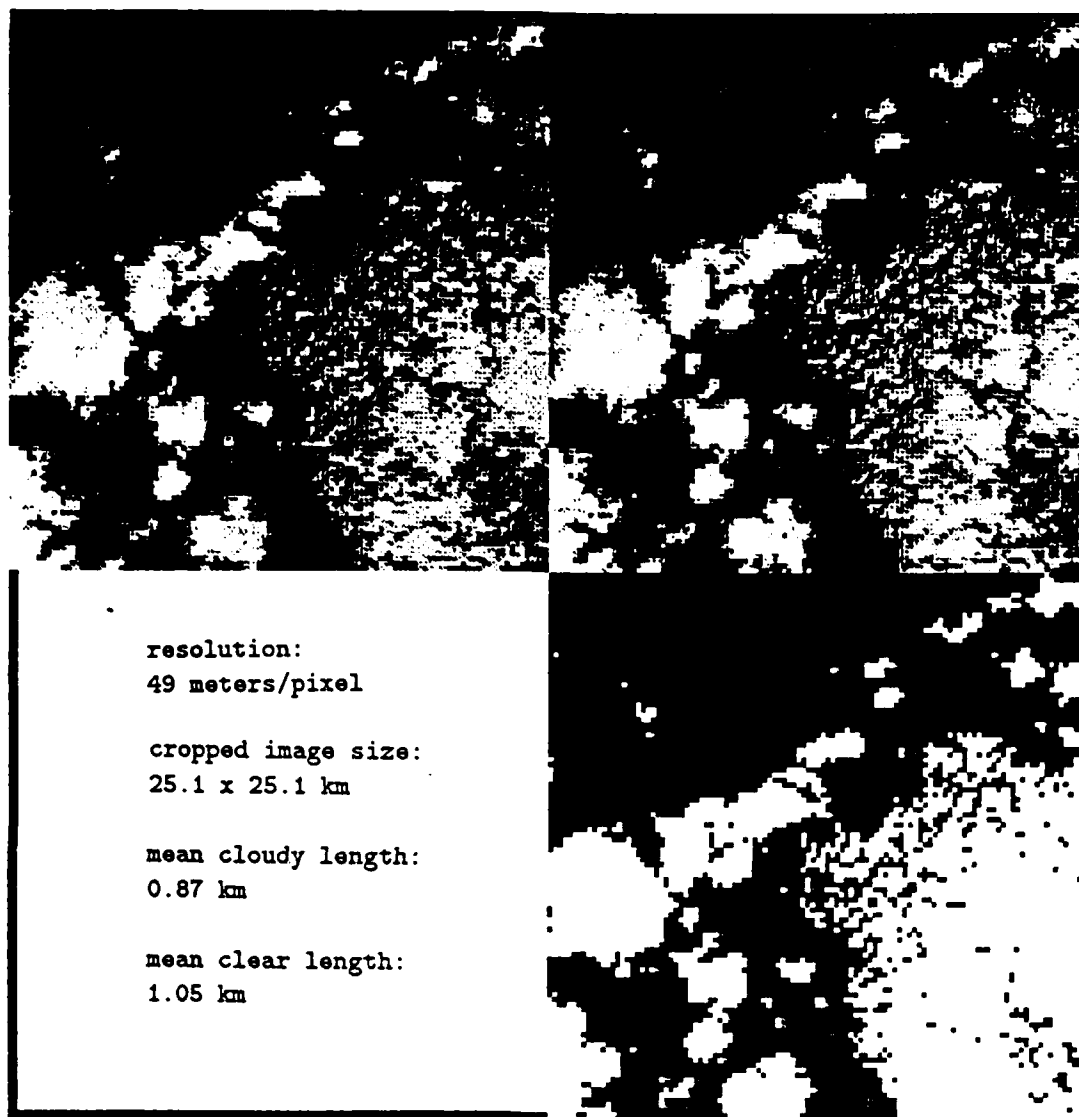
Figure A.4   TREATMENTS AND STATISTICS OF IMAGE 4
(See Appendix A text for figure explanation.)

resolution:
49 meters/pixel

cropped image size:
25.1 x 25.1 km

mean cloudy length:
1.15 km

mean clear length:
0.83 km
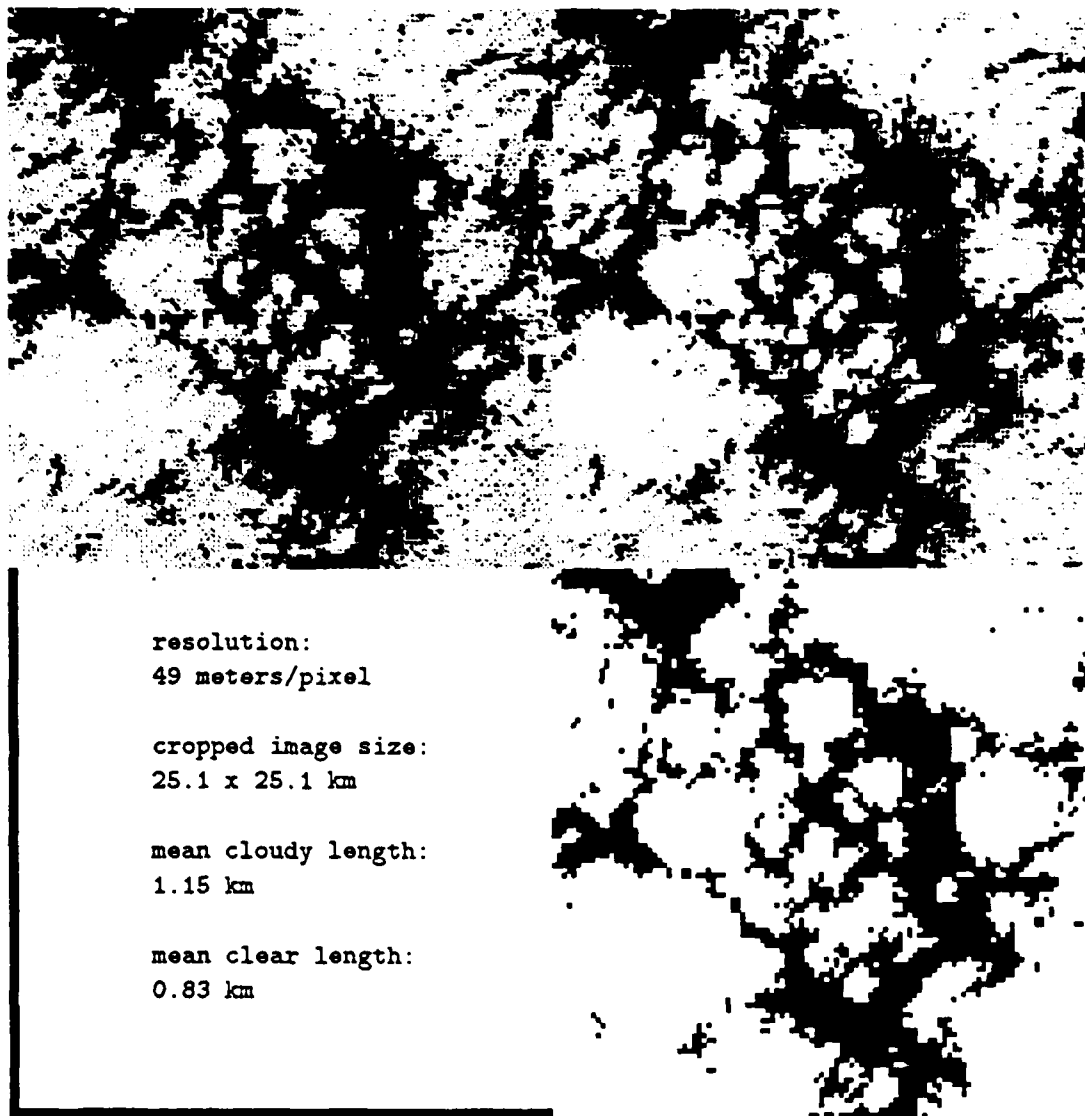
Figure A.5   TREATMENTS AND STATISTICS OF IMAGE 5
(See Appendix A text for figure explanation.)

resolution:
49 meters/pixel

cropped image size:
25.1 x 25.1 km

mean cloudy length:
0.51 km

mean clear length:
0.54 km
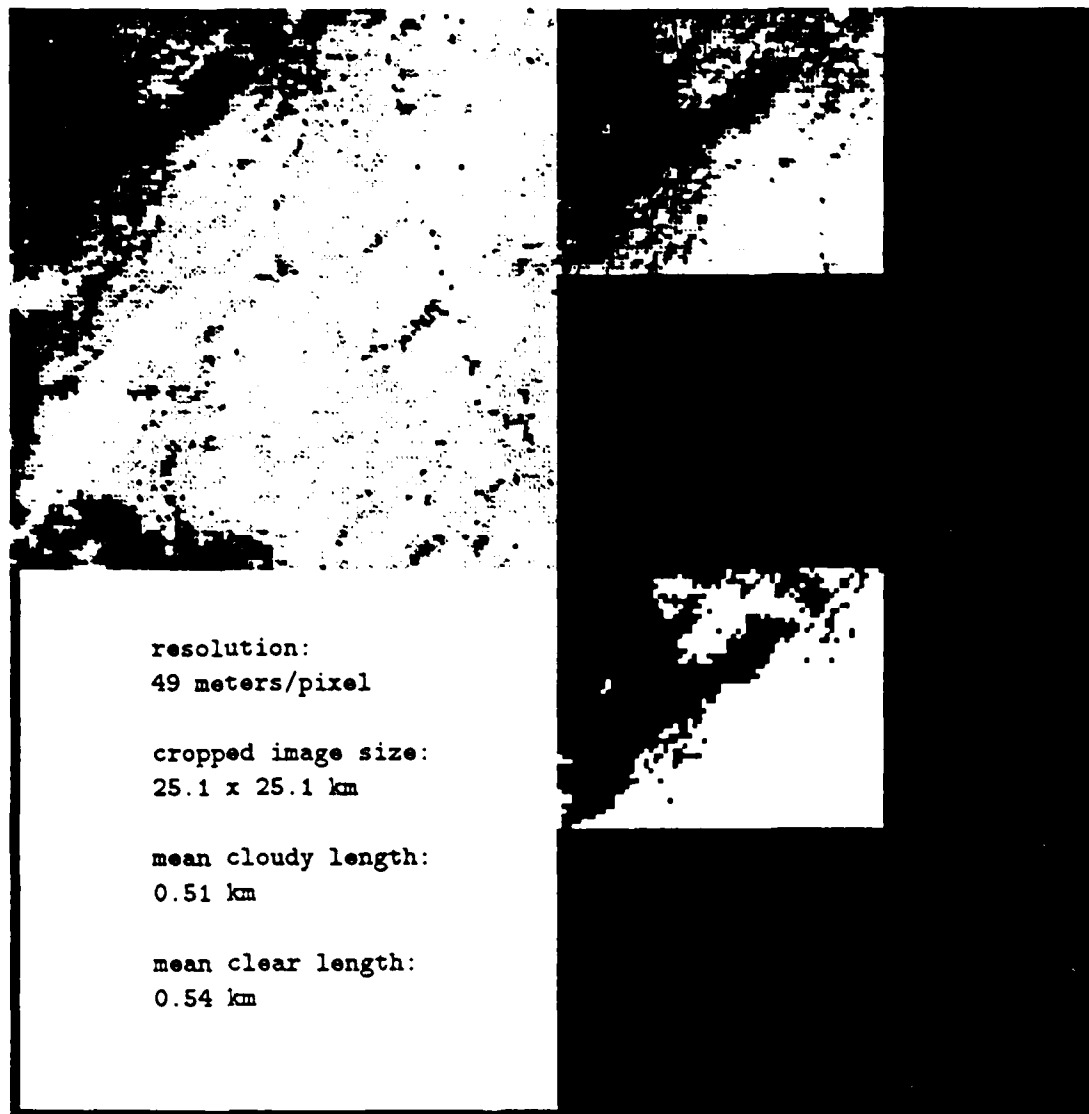
Figure A.6   TREATMENTS AND STATISTICS OF IMAGE 6
(See Appendix A text for figure explanation.)

```
resolution:
49 meters/pixel

cropped image size:
25.1 x 25.1 km

mean cloudy length:
1.08 km

mean clear length:
1.08 km
```
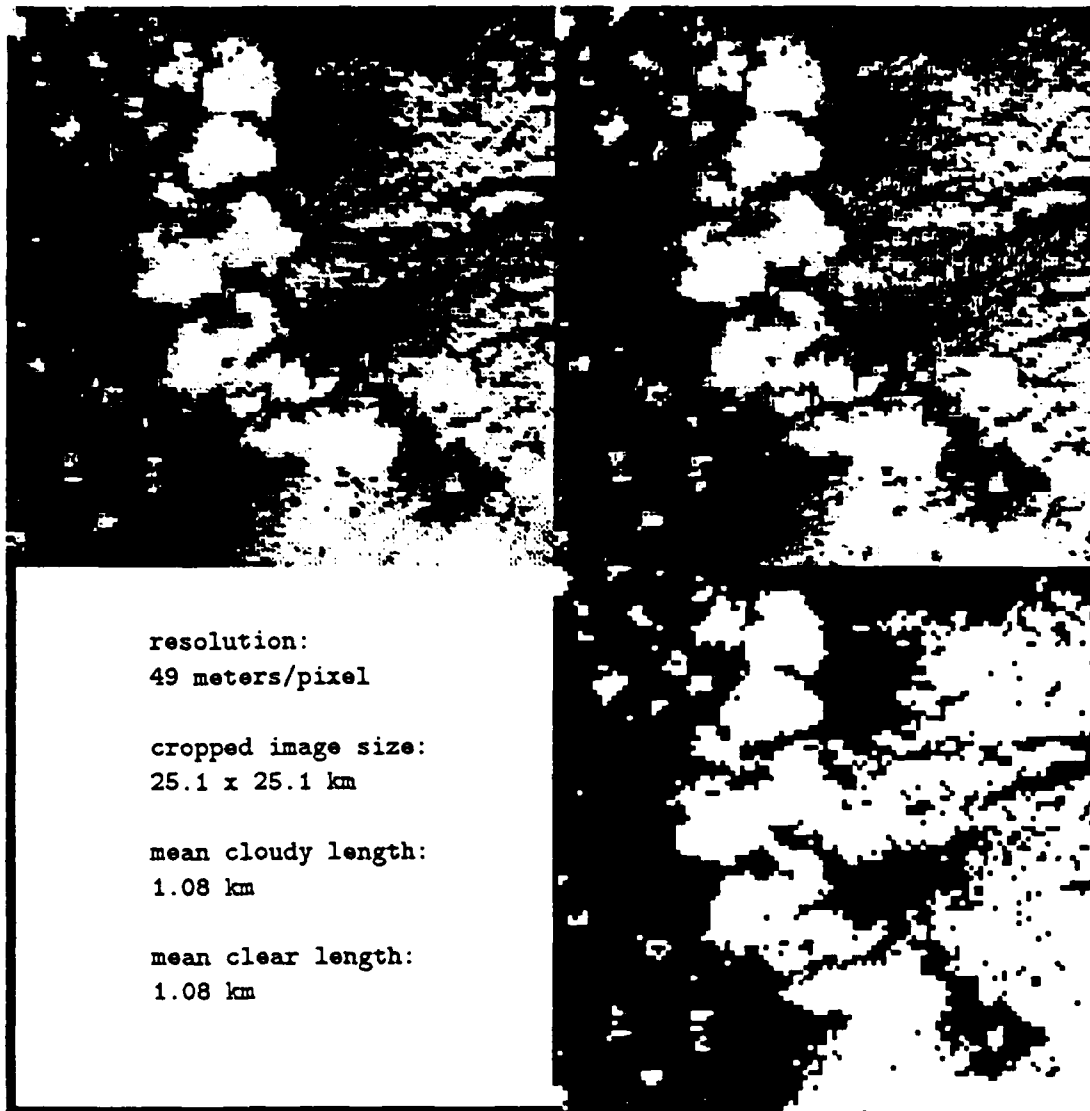
Figure A.7   TREATMENTS AND STATISTICS OF IMAGE 7
(See Appendix A text for figure explanation.)

resolution:
49 meters/pixel

cropped image size:
25.1 x 25.1 km

mean cloudy length:
0.80 km

mean clear length:
1.45 km
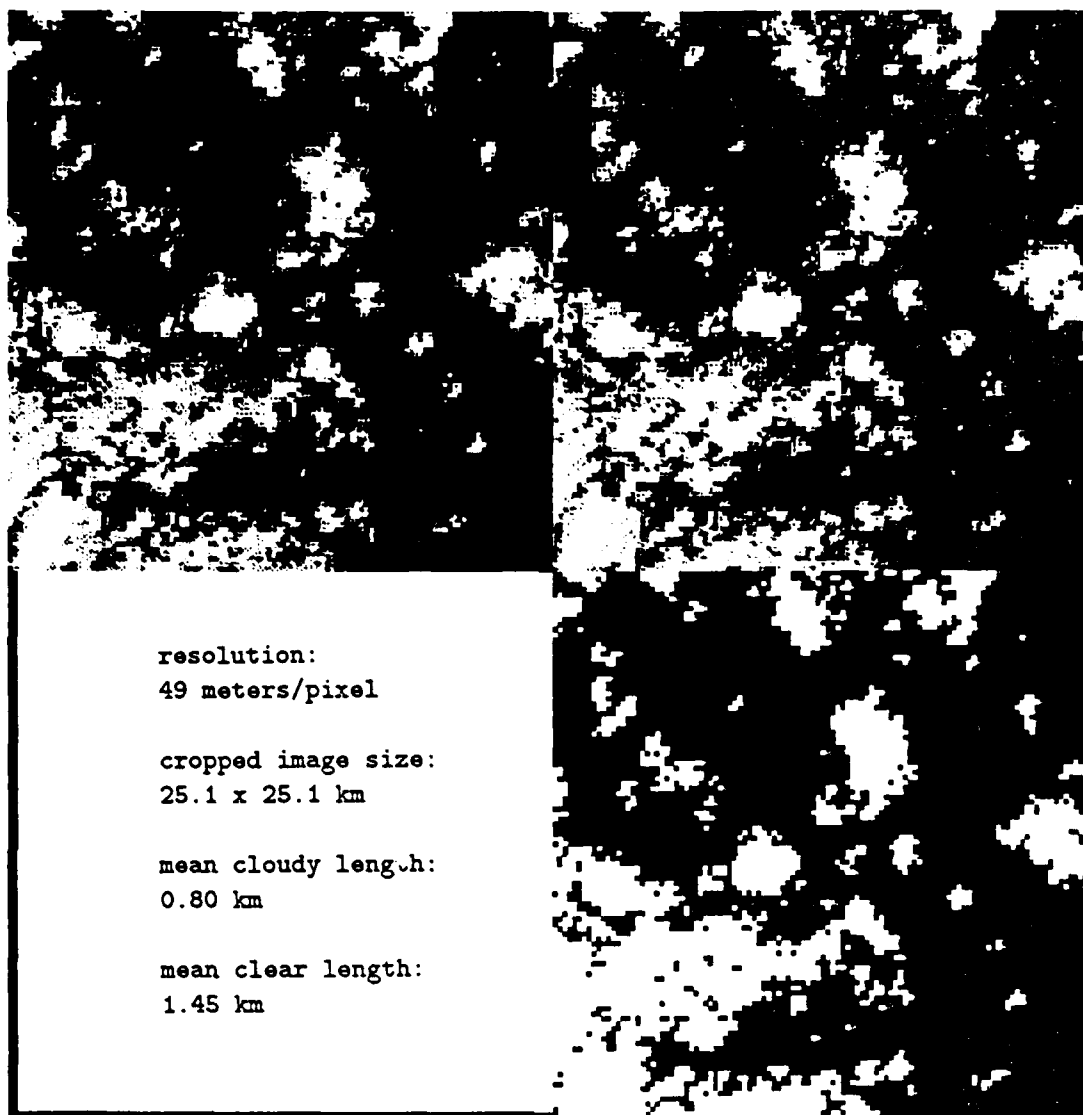
Figure A.8   TREATMENTS AND STATISTICS OF IMAGE 8
(See Appendix A text for figure explanation.)

```
resolution:
99 meters/pixel

cropped image size:
50.7 x 50.7 km

mean cloudy length:
2.02 km

mean clear length:
1.56 km
```
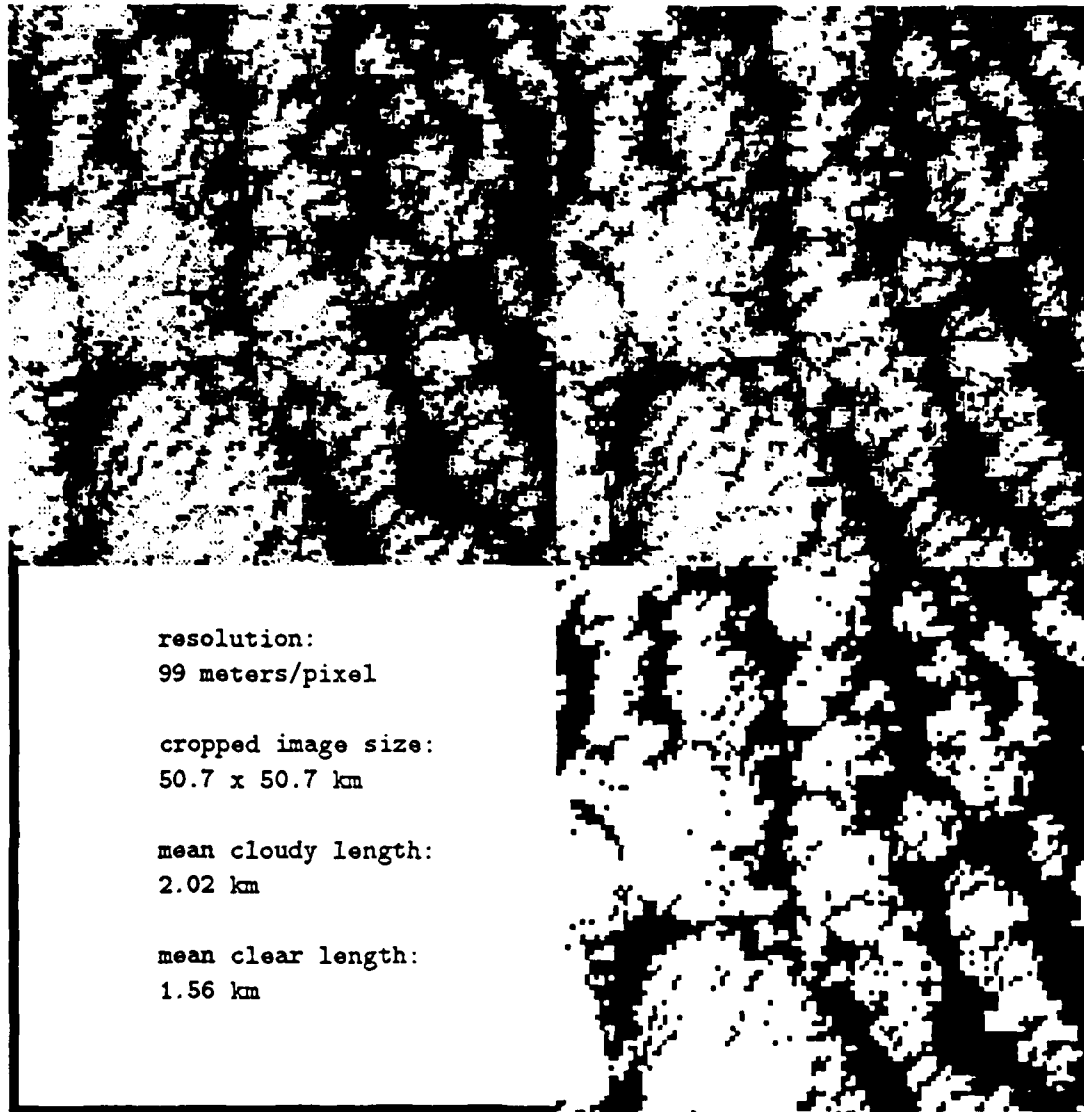
Figure A.9   TREATMENTS AND STATISTICS OF IMAGE 9
(See Appendix A text for figure explanation.)

resolution:
99 meters/pixel

cropped image size:
50.7 x 50.7 km

mean cloudy length:
1.39 km

mean clear length:
4.17 km
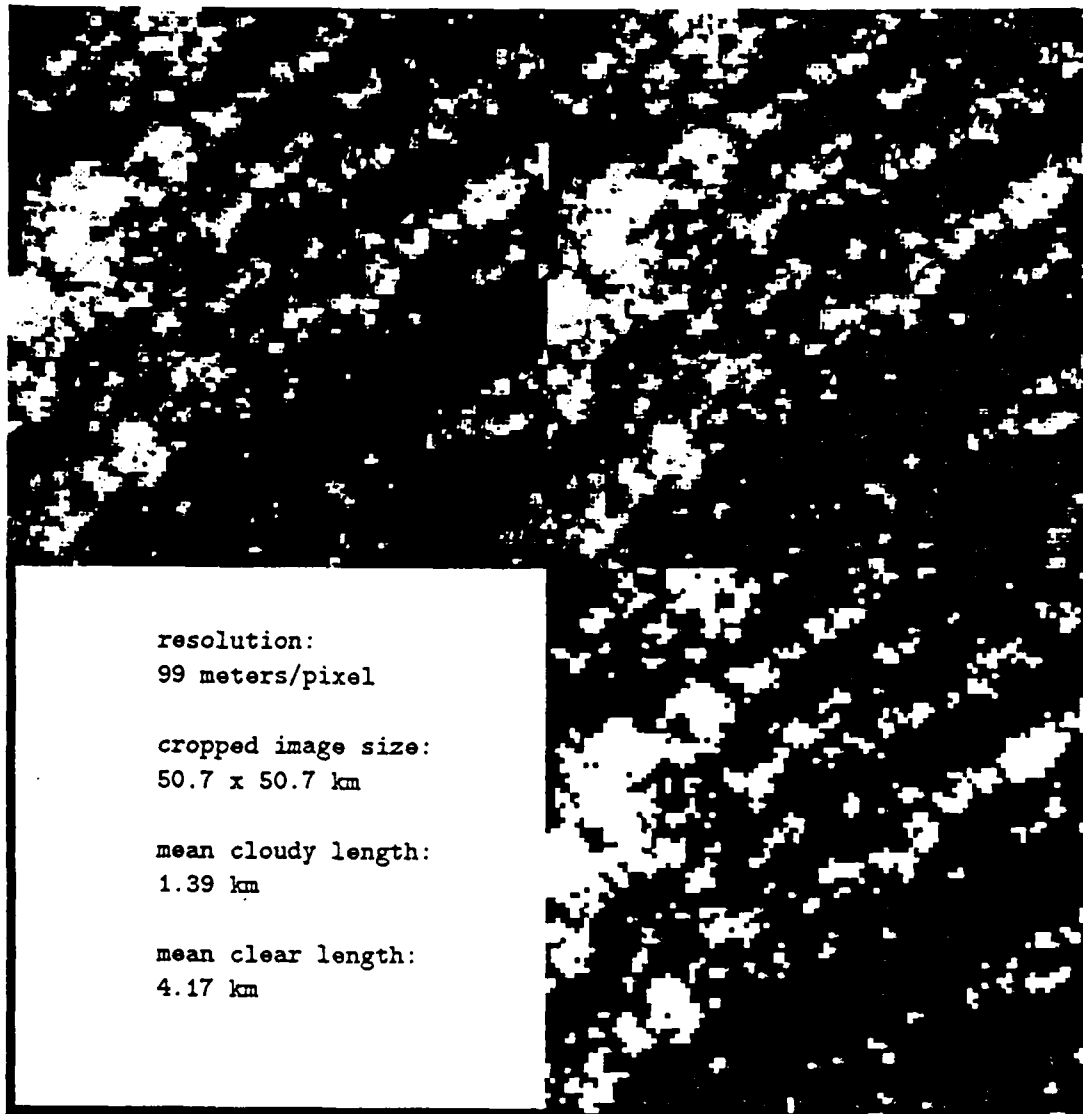
Figure A.10   TREATMENTS AND STATISTICS OF IMAGE 10
(See Appendix A text for figure explanation.)

resolution:
99 meters/pixel

cropped image size:
50.7 x 50.7 km

mean cloudy length:
1.76 km

mean clear length:
3.99 km
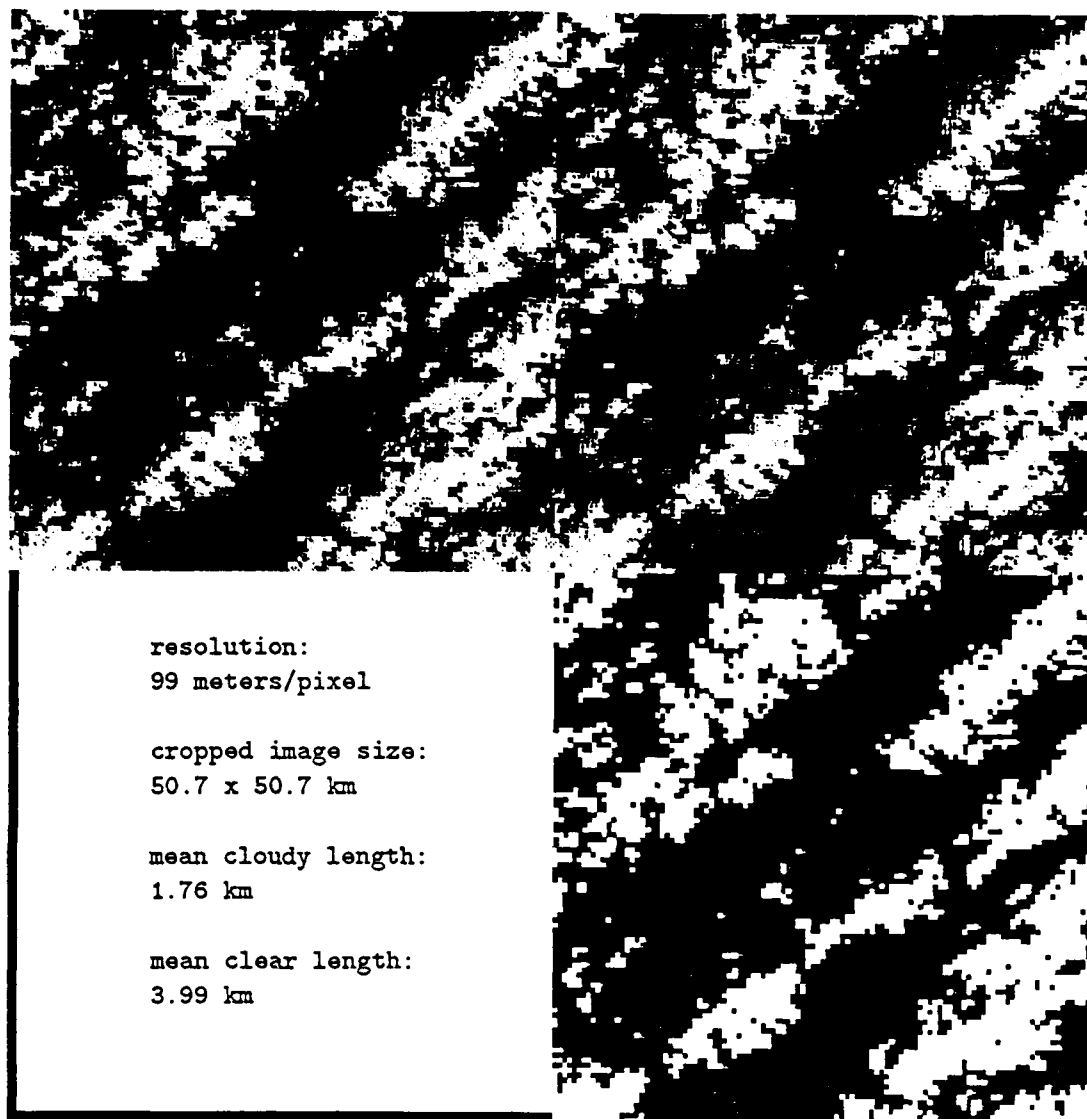
Figure A.11 TREATMENTS AND STATISTICS OF IMAGE 11
(See Appendix A text for figure explanation.)

resolution:
99 meters/pixel

cropped image size:
50.7 x 50.7 km

mean cloudy length:
1.23 km

mean clear length:
2.13 km
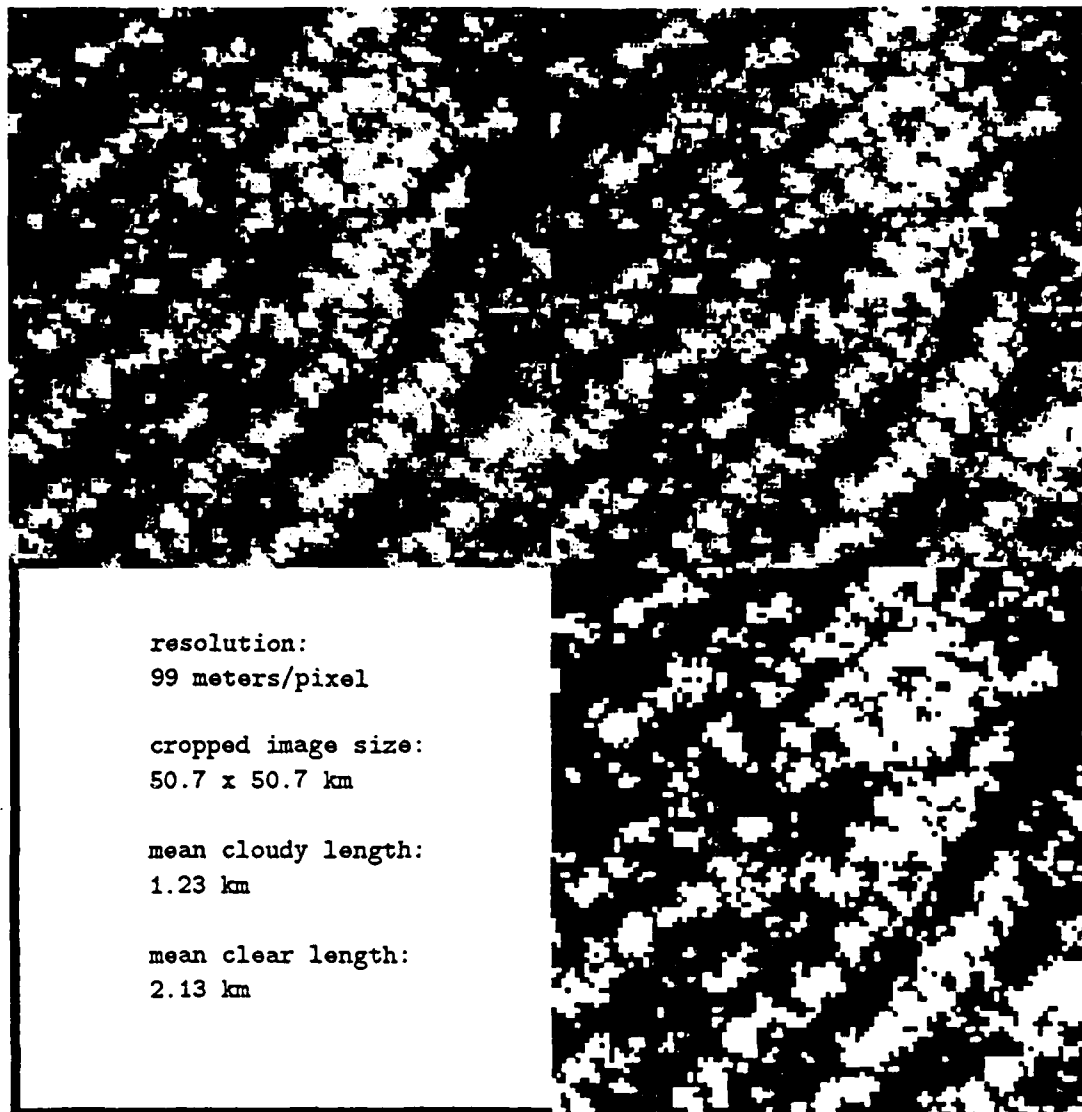
Figure A.12 TREATMENTS AND STATISTICS OF IMAGE 12
(See Appendix A text for figure explanation.)

resolution:
300 meters/pixel

cropped image size:
72.3 x 72.0 km

mean cloudy length:
2.55 km

mean clear length:
2.11 km
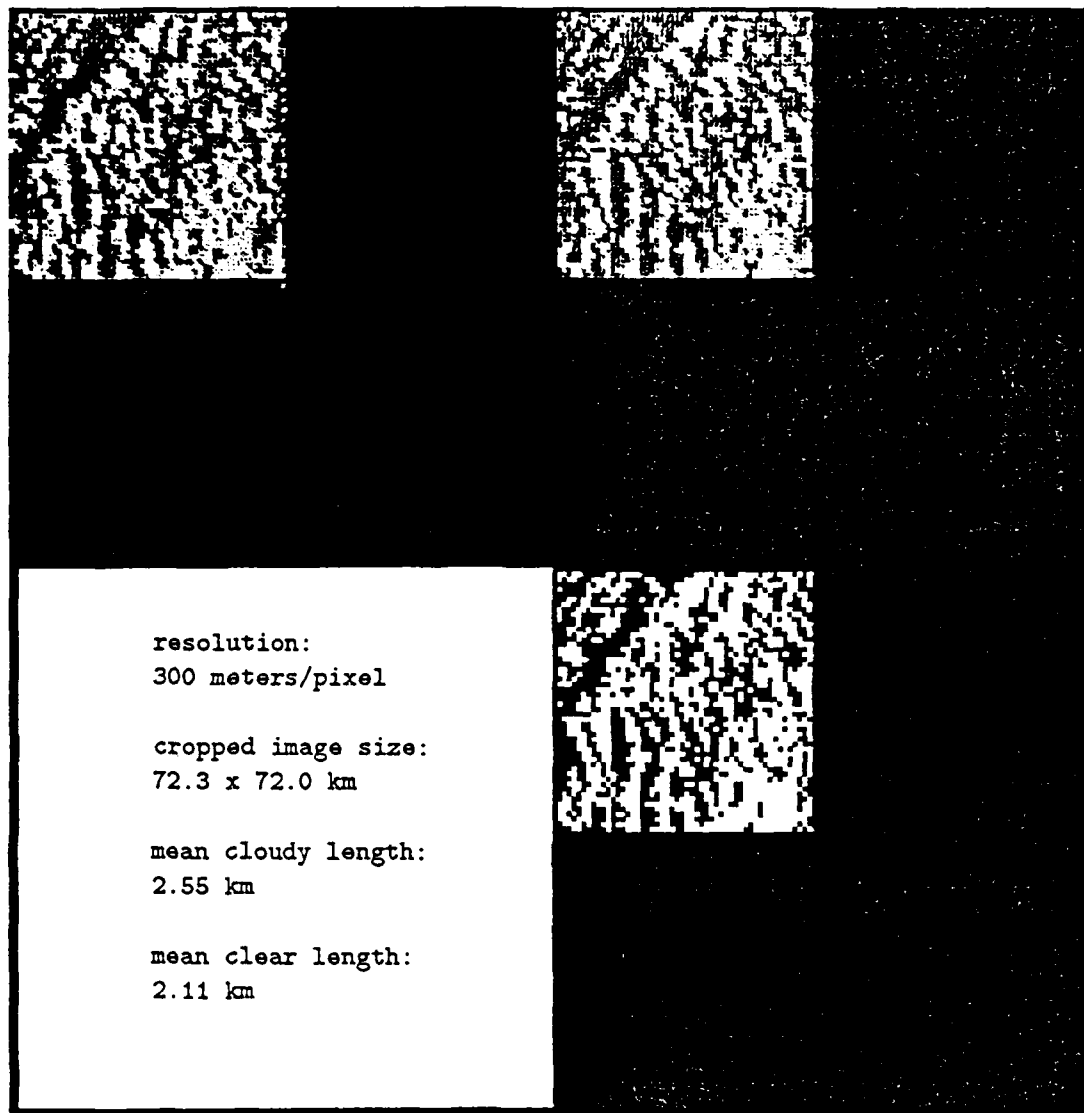
Figure A.13 TREATMENTS AND STATISTICS OF IMAGE 13
(See Appendix A text for figure explanation.)

resolution:
225 meters/pixel

cropped image size:
54.2 x 54.0 km

mean cloudy length:
2.83 km

mean clear length:
1.14 km
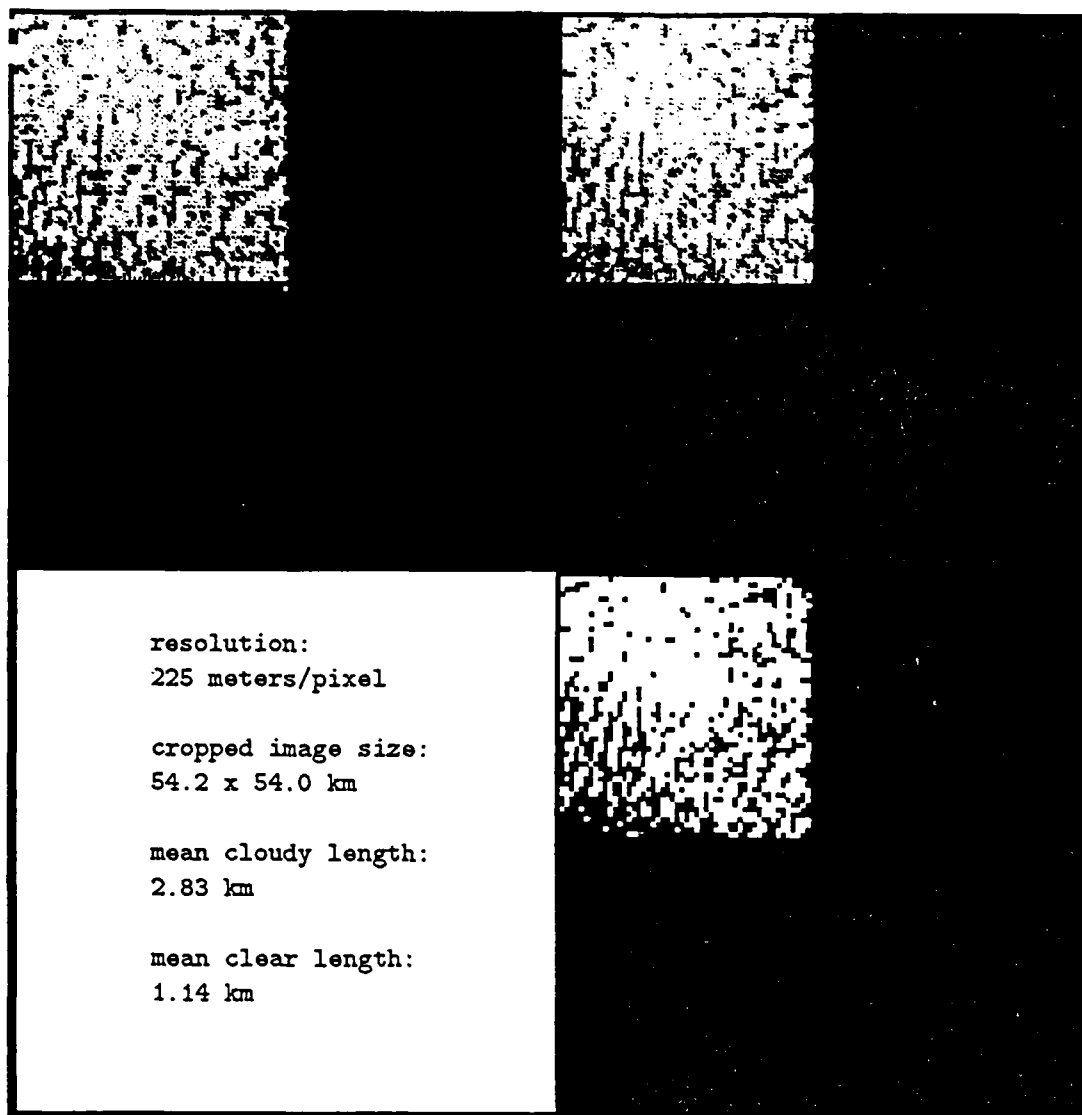
Figure A.14   TREATMENTS AND STATISTICS OF IMAGE 14
(See Appendix A text for figure explanation.)

resolution:
65 meters/pixel

cropped image size:
16.9 x 23.7 km

mean cloudy length:
0.96 km

mean clear length:
0.67 km
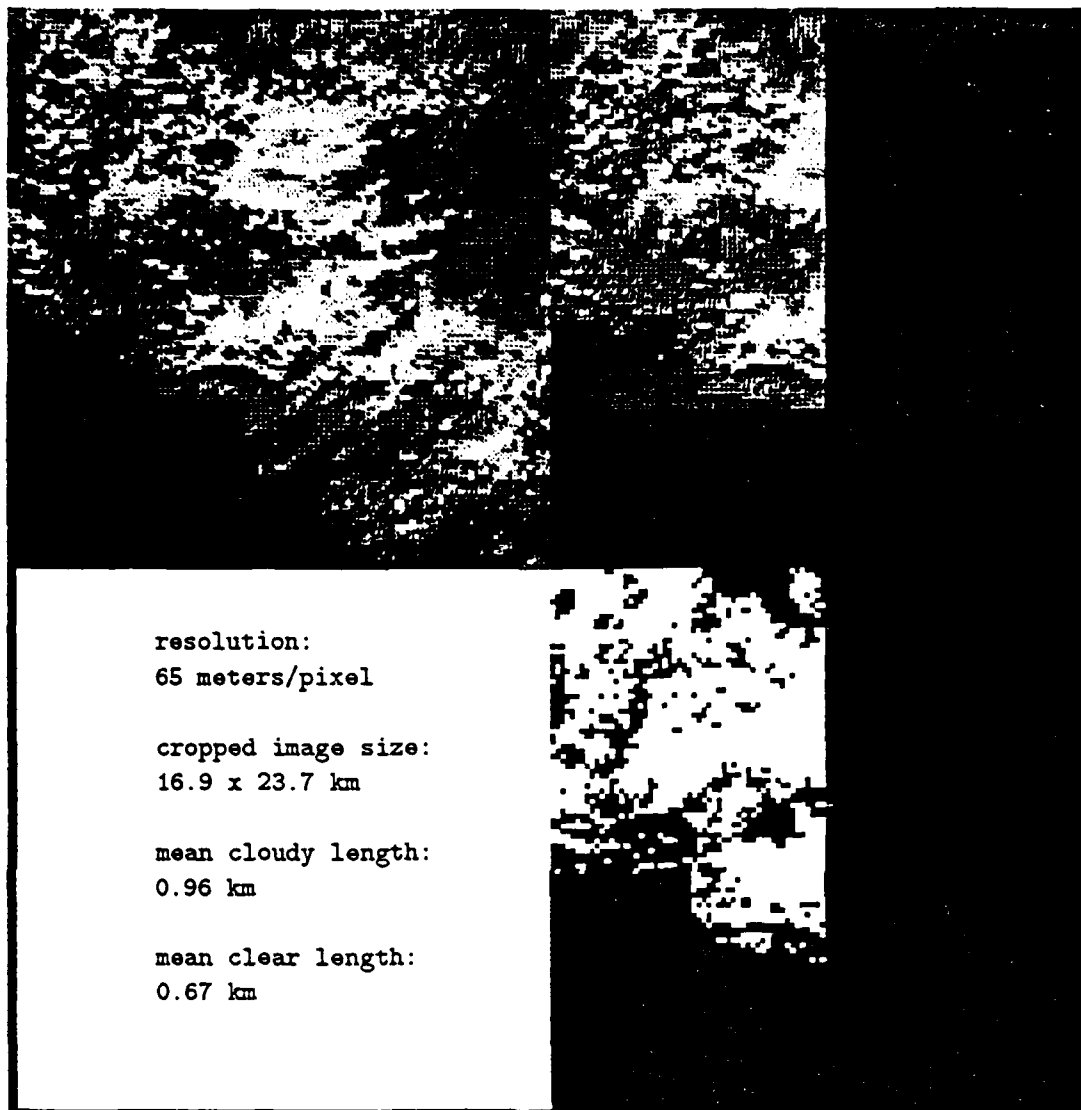
Figure A.15 TREATMENTS AND STATISTICS OF IMAGE 15
(See Appendix A text for figure explanation.)

resolution:
58 meters/pixel

cropped image size:
26.9 x 21.6 km

mean cloudy length:
0.43 km

mean clear length:
0.83 km
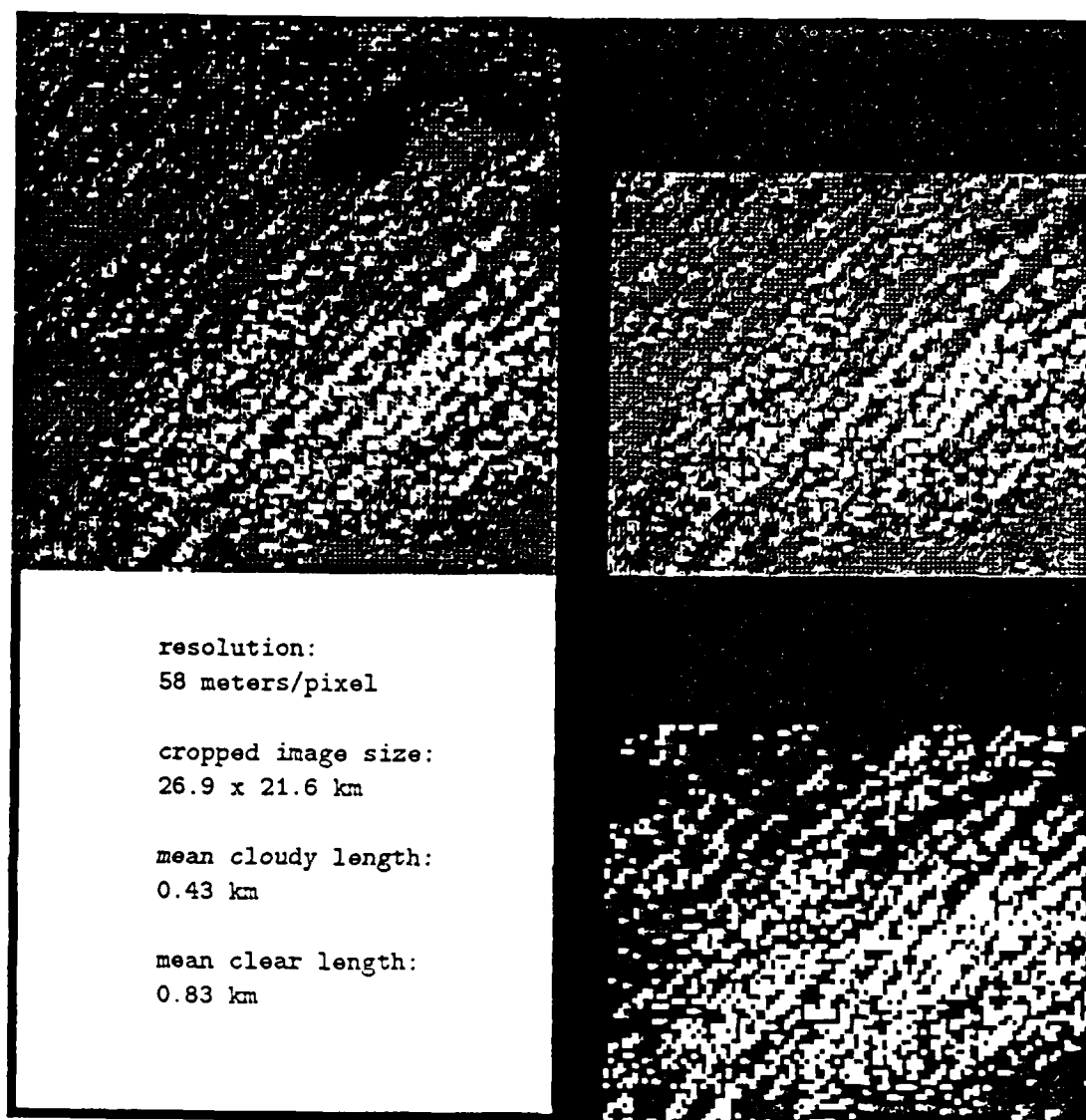
Figure A.16 TREATMENTS AND STATISTICS OF IMAGE 16
(See Appendix A text for figure explanation.)

APPENDIX B

STATISTICAL TREATMENT:

The intent of this research was to successfully model cloudy and clear interval length PDFs and 1-CDFs. Before prospective models could be treated, a statistical analysis method had to be chosen. The method would be used to perform regressions and to determine the goodness-of-fit of the models to the data.

It is desirable that any statistical analysis method be both metric and congruent. As pointed out by Mielke (1986), a method is metric if (i) $D_{i,j} \geq 0$ and $D_{i,i} = 0$, (ii) $D_{i,j} = D_{j,i}$ (symmetry), and (iii) $D_{i,j} + D_{j,k} \geq D_{i,k}$ (triangle inequality). The first conditions state that the distance between two points must be greater than or equal to zero and that the distance from a point to itself is zero. The second condition states that the distance measured between two points is the same regardless of which point is considered first. The third condition states that the sum of any two sides of a triangle must be greater than or equal to the measure of the third. These conditions, if met, aid in the definition of a Euclidean space.

The domain of observations between which comparisons are made (the data space) is ususally perceived as a Euclidean space and is, therefore, metric. It is ideal to attain congruence between the data space and analysis space. This requires that the analysis space be metric as well. This often goes unrecognized as evidenced by the frequent use of least squares analysis methods. The analysis space of these methods is incongruent in that it does not match any Euclidean data space and non-metric because it does not satisfy the triangle inequality. Least squares methods also require that the variance be homogeneous and normally distributed about the mean. For all these reasons, least squares methods were abandoned.

There are statistical analysis methods that satisfy the metric and congruent require-ments. They are based on least absolute deviations rather than least squared deviations.

Treatment of the data space using these methods is in keeping with our intuitive Euclidean perceptions. Least absolute methods were used in Chapter 3.

APPENDIX C

MEAN ERRORS FOR INDIVIDUAL IMAGES:

These tables present G values (mean errors) for every condition/case combination for every image. The exponential model is outperformed in only 3 of 64 instances. In these instances the hyperbolic model accurately treats the CLEAR/1-CDF CONDITION/CASE.

TABLE C.1   MEAN ERRORS OF MODELED PROBABILITIES FOR IMAGE 1

CONDITION	CASE	MODEL	TOTAL ERROR	NUMBER OF OBSERVATIONS	G^* (mean error)
CLOUDY	Pr(X=x) OR PDF	exp	.3694	104	.0036**
		hyp	2.0435	104	.0196
		log	1.0015	104	.0096
	Pr(X≥x) OR 1-CDF	exp	2.0453	179	.0114**
		hyp	3.2768	179	.0183
		log	5.1769	179	.0289
CLEAR	Pr(X=x) OR PDF	exp	.2897	41	.0071**
		hyp	1.8508	41	.0451
		log	1.0051	41	.0245
	Pr(X≥x) OR 1-CDF	exp	1.0330	89	.0116**
		hyp	2.5703	89	.0289
		log	4.5366	89	.0510

* Multiply these values by 100 to find the percent error.
** These models performed best in their respective CONDITION/CASE combinations.

TABLE C.2   MEAN ERRORS OF MODELED PROBABILITIES FOR IMAGE 2

CONDITION	CASE	MODEL	TOTAL ERROR	NUMBER OF OBSERVATIONS	G^* (mean error)
CLOUDY	Pr(X=x) OR PDF	exp	.3745	110	.0034**
		hyp	1.3321	110	.0121
		log	1.0534	110	.0096
	Pr(X≥x) OR 1-CDF	exp	5.8057	137	.0424**
		hyp	9.4241	137	.0688
		log	16.7178	137	.1220
CLEAR	Pr(X=x) OR PDF	exp	.4491	133	.0034**
		hyp	1.3543	133	.0102
		log	1.0351	133	.0078
	Pr(X≥x) OR 1-CDF	exp	11.3177	218	.0519**
		hyp	13.2375	218	.0607
		log	22.0372	218	.1011

*   Multiply these values by 100 to find the percent error.

** These models performed best in their respective
     CONDITION/CASE combinations.

TABLE C.3   MEAN ERRORS OF MODELED PROBABILITIES FOR IMAGE 3

CONDITION	CASE	MODEL	TOTAL ERROR	NUMBER OF OBSERVATIONS	G^* (mean error)
CLOUDY	Pr(X=x) OR PDF	exp	.2865	72	.0040**
		hyp	1.6511	72	.0229
		log	1.0167	72	.0141
	Pr(X≥x) OR 1-CDF	exp	1.5171	104	.0146**
		hyp	8.5522	104	.0822
		log	12.5533	104	.1207
CLEAR	Pr(X=x) OR PDF	exp	6500	241	.0027**
		hyp	1.1510	241	.0048
		log	1.0471	241	.0043
	Pr(X≥x) OR 1-CDF	exp	26.3525	418	.0630**
		hyp	27.3945	418	.0655
		log	43.0012	418	.1029

*   Multiply these values by 100 to find the percent error.
** These models performed best in their respective CONDITION/CASE combinations.

TABLE C.4   MEAN ERRORS OF MODELED PROBABILITIES FOR IMAGE 4

CONDITION	CASE	MODEL	TOTAL ERROR	NUMBER OF OBSERVATIONS	G^* (mean error)
CLOUDY	Pr(X=x) OR PDF	exp	.5232	134	.0039**
		hyp	1.1601	134	.0089
		log	1.0401	134	.0078
	Pr(X≥x) OR 1-CDF	exp	7.6218	173	.0441**
		hyp	8.6531	173	.0500
		log	16.1890	173	.0936
CLEAR	Pr(X=x) OR PDF	exp	.8893	167	.0053**
		hyp	1.4772	167	.0088
		log	1.0023	167	.0060
	Pr(X≥x) OR 1-CDF	exp	26.1830	413	.0634**
		hyp	16.0812	413	.0389
		log	20.7773	413	.0503

* Multiply these values by 100 to find the percent error.

** These models performed best in their respective CONDITION/CASE combinations.

TABLE C.5   MEAN ERRORS OF MODELED PROBABILITIES FOR IMAGE 5

CONDITION	CASE	MODEL	TOTAL ERROR	NUMBER OF OBSERVATIONS	G^* (mean error)
CLOUDY	Pr(X=x) OR PDF	exp	.4661	163	.0029**
		hyp	1.5283	163	.0094
		log	1.0091	163	.0062
	Pr(X≥x) OR 1-CDF	exp	8.1467	269	.0303**
		hyp	17.4624	269	.0649
		log	22.9170	269	.0852
CLEAR	Pr(X=x) OR PDF	exp	.5476	121	.0045**
		hyp	1.0653	121	.0088
		log	1.0433	121	.0086
	Pr(X≥x) OR 1-CDF	exp	5.6682	141	.0402**
		hyp	8.9198	141	.0633
		log	15.6746	141	.1112

*   Multiply these values by 100 to find the percent error.

** These models performed best in their respective CONDITION/CASE combinations.

TABLE C.6   MEAN ERRORS OF MODELED PROBABILITIES FOR IMAGE 6

CONDITION	CASE	MODEL	TOTAL ERROR	NUMBER OF OBSERVATIONS	G^* (mean error)
CLOUDY	Pr(X=x) OR PDF	exp	.4193	54	.0078**
		hyp	1.3693	54	.0254
		log	1.0286	54	.0190
	Pr(X≥x) OR 1-CDF	exp	1.9194	77	.0249**
		hyp	5.7590	77	.0748
		log	9.8395	77	.1278
CLEAR	Pr(X=x) OR PDF	exp	.6537	63	.0104**
		hyp	1.3181	63	.0209
		log	1.0157	63	.0161
	Pr(X≥x) OR 1-CDF	exp	5.9228	114	.0520**
		hyp	6.2155	114	.0545
		log	10.3621	114	.0909

* Multiply these values by 100 to find the percent error.
** These models performed best in their respective
   CONDITION/CASE combinations.

TABLE C.7   MEAN ERRORS OF MODELED PROBABILITIES FOR IMAGE 7

CONDITION	CASE	MODEL	TOTAL ERROR	NUMBER OF OBSERVATIONS	G^* (mean error)
CLOUDY	Pr(X=x) OR PDF	exp	.5877	161	.0037**
		hyp	1.2357	161	.0077
		log	1.0163	161	.0063
	Pr(X≥x) OR 1-CDF	exp	9.1461	215	.0425**
		hyp	13.0889	215	.0609
		log	20.8564	215	.0970
CLEAR	Pr(X=x) OR PDF	exp	.7761	181	.0043**
		hyp	1.1291	181	.0062
		log	1.0165	181	.0056
	Pr(X≥x) OR 1-CDF	exp	15.3357	287	.0534
		hyp	13.5228	287	.0471**
		log	20.7489	287	.0723

* Multiply these values by 100 to find the percent error.

** These models performed best in their respective CONDITION/CASE combinations.

TABLE C.8   MEAN ERRORS OF MODELED PROBABILITIES FOR IMAGE 8

CONDITION	CASE	MODEL	TOTAL ERROR	NUMBER OF OBSERVATIONS	G^* (mean error)
CLOUDY	Pr(X=x) OR PDF	exp	.5543	127	.0044**
		hyp	1.3901	127	.0109
		log	1.0121	127	.0080
	Pr(X≥x) OR 1-CDF	exp	9.6981	216	.0449**
		hyp	10.3378	216	.0479
		log	15.7208	216	.0728
CLEAR	Pr(X=x) OR PDF	exp	.7637	200	.0038**
		hyp	.9538	200	.0048
		log	1.0417	200	.0052
	Pr(X≥x) OR 1-CDF	exp	15.8131	275	.0575**
		hyp	15.8882	275	.0578
		log	26.4392	275	.0961

* Multiply these values by 100 to find the percent error.

** These models performed best in their respective CONDITION/CASE combinations.

TABLE C.9   MEAN ERRORS OF MODELED PROBABILITIES FOR IMAGE 9

CONDITION	CASE	MODEL	TOTAL ERROR	NUMBER OF OBSERVATIONS	G^* (mean error)
CLOUDY	Pr(X=x) OR PDF	exp	.4239	159	.0027**
		hyp	1.3984	159	.0088
		log	1.0241	159	.0064
	Pr(X≥x) OR 1-CDF	exp	6.9153	193	.0358**
		hyp	12.5797	193	.0652
		log	19.0751	193	.0988
CLEAR	Pr(X=x) OR PDF	exp	.5087	112	.0045**
		hyp	1.5275	112	.0136
		log	1.0090	112	.0090
	Pr(X≥x) OR 1-CDF	exp	3.6541	159	.0230**
		hyp	11.4825	159	.0722
		log	15.5222	159	.0976

* Multiply these values by 100 to find the percent error.

** These models performed best in their respective CONDITION/CASE combinations.

TABLE C.10   MEAN ERRORS OF MODELED PROBABILITIES FOR IMAGE 10

CONDITION	CASE	MODEL	TOTAL ERROR	NUMBER OF OBSERVATIONS	G^* (mean error)
CLOUDY	Pr(X=x) OR PDF	exp	.3393	91	.0037**
		hyp	1.6636	91	.0183
		log	1.0156	91	.0112
	Pr(X≥x) OR 1-CDF	exp	4.2427	132	.0321**
		hyp	9.0050	132	.0682
		log	13.4479	132	.1019
CLEAR	Pr(X=x) OR PDF	exp	.4223	223	.0019**
		hyp	1.4926	223	.0067
		log	1.0215	223	.0046
	Pr(X≥x) OR 1-CDF	exp	15.2218	418	.0364**
		hyp	29.4698	418	.0705
		log	40.1674	418	.0961

* Multiply these values by 100 to find the percent error.

** These models performed best in their respective CONDITION/CASE combinations.

TABLE C.11   MEAN ERRORS OF MODELED PROBABILITIES FOR IMAGE 11

CONDITION	CASE	MODEL	TOTAL ERROR	NUMBER OF OBSERVATIONS	G^* (mean error)
CLOUDY	Pr(X=x) OR PDF	exp	.2036	94	.0022**
		hyp	1.5248	94	.0162
		log	1.0343	94	.0110
	Pr(X≥x) OR 1-CDF	exp	2.3883	115	.0208**
		hyp	11.8123	115	.1027
		log	17.2083	115	.1496
CLEAR	Pr(X=x) OR PDF	exp	.7210	229	.0031**
		hyp	1.2036	229	.0053
		log	1.0305	229	.0045
	Pr(X≥x) OR 1-CDF	exp	14.6087	352	.0415**
		hyp	26.5452	352	.0754
		log	37.6494	352	.1070

* Multiply these values by 100 to find the percent error.

** These models performed best in their respective CONDITION/CASE combinations.

TABLE C.12   MEAN ERRORS OF MODELED PROBABILITIES FOR IMAGE 12

CONDITION	CASE	MODEL	TOTAL ERROR	NUMBER OF OBSERVATIONS	G^* (mean error)
CLOUDY	Pr(X=x) OR PDF	exp	.2148	83	.0026**
		hyp	1.4919	83	.0180
		log	1.0263	83	.0124
	Pr(X≥x) OR 1-CDF	exp	.7901	90	.0088**
		hyp	7.7721	90	.0864
		log	11.9415	90	.1327
CLEAR	Pr(X=x) OR PDF	exp	.4609	150	.0031**
		hyp	1.3890	150	.0093
		log	1.0212	150	.0068
	Pr(X≥x) OR 1-CDF	exp	4.5985	178	.0258**
		hyp	14.6900	178	.0825
		log	20.5540	178	.1155

* Multiply these values by 100 to find the percent error.

** These models performed best in their respective CONDITION/CASE combinations.

TABLE C.13   MEAN ERRORS OF MODELED PROBABILITIES FOR IMAGE 13

CONDITION	CASE	MODEL	TOTAL ERROR	NUMBER OF OBSERVATIONS	G^* (mean error)
CLOUDY	Pr(X=x) OR PDF	exp	.1987	45	.0044**
		hyp	1.7347	45	.0385
		log	1.0242	45	.0228
	Pr(X≥x) OR 1-CDF	exp	1.5423	50	.0308**
		hyp	5.3337	50	.1067
		log	8.2453	50	.1649
CLEAR	Pr(X=x) OR PDF	exp	.1094	42	.0026**
		hyp	1.8956	42	.0451
		log	1.0082	42	.0240
	Pr(X≥x) OR 1-CDF	exp	.8236	58	.0142**
		hyp	4.5904	58	.0791
		log	6.9515	58	.1199

*   Multiply these values by 100 to find the percent error.
** These models performed best in their respective
     CONDITION/CASE combinations.

TABLE C.14   MEAN ERRORS OF MODELED PROBABILITIES FOR IMAGE 14

CONDITION	CASE	MODEL	TOTAL ERROR	NUMBER OF OBSERVATIONS	G^* (mean error)
CLOUDY	Pr(X=x) OR PDF	exp	.5340	104	.0051**
		hyp	1.6078	104	.0155
		log	1.0092	104	.0097
	Pr(X≥x) OR 1-CDF	exp	7.6293	179	.0426**
		hyp	7.8503	179	.0436
		log	12.2218	179	.0683
CLEAR	Pr(X=x) OR PDF	exp	.4471	41	.0109**
		hyp	2.1281	41	.0519
		log	1.0003	41	.0244
	Pr(X≥x) OR 1-CDF	exp	2.7194	89	.0306**
		hyp	3.1798	89	.0357
		log	5.0265	89	.0555

*   Multiply these values by 100 to find the percent error.

**   These models performed best in their respective
     CONDITION/CASE combinations.

TABLE C.15   MEAN ERRORS OF MODELED PROBABILITIES FOR IMAGE 15

CONDITION	CASE	MODEL	TOTAL ERROR	NUMBER OF OBSERVATIONS	G^* (mean error)
CLOUDY	Pr(X=x) OR PDF	exp	.4189	91	.0046**
		hyp	1.3890	91	.0153
		log	1.0282	91	.0089
	Pr(X≥x) OR 1-CDF	exp	3.3681	116	.0290**
		hyp	8.7184	116	.0075
		log	13.7842	116	.1188
CLEAR	Pr(X=x) OR PDF	exp	.4187	76	.0055**
		hyp	1.3934	76	.0183
		log	1.0180	76	.0134
	Pr(X≥x) OR 1-CDF	exp	4.2276	105	.0403**
		hyp	5.6749	105	.0540
		log	9.7626	105	.0930

* Multiply these values by 100 to find the percent error.

** These models performed best in their respective CONDITION/CASE combinations.

TABLE C.16   MEAN ERRORS OF MODELED PROBABILITIES FOR IMAGE 16

CONDITION	CASE	MODEL	TOTAL ERROR	NUMBER OF OBSERVATIONS	G^* (mean error)
CLOUDY	Pr(X=x) OR PDF	exp	.2351	48	.0049**
		hyp	2.3885	48	.0498
		log	1.0024	48	.0209
	Pr(X≥x) OR 1-CDF	exp	1.3048	66	.0198**
		hyp	5.4727	66	.0829
		log	7.4080	66	.1122
CLEAR	Pr(X=x) OR PDF	exp	.5597	134	.0042**
		hyp	1.8827	134	.0141
		log	1.0035	134	.0075
	Pr(X≥x) OR 1-CDF	exp	11.8303	254	.0466
		hyp	10.2144	254	.0402**
		log	14.0700	254	.0554

* Multiply these values by 100 to find the percent error.

** These models performed best in their respective CONDITION/CASE combinations.

APPENDIX D

EXTREMELY HIGH RESOLUTION IMAGERY FROM THE SPACE SHUTTLE ORBITER:

Imagery of the cloudy atmosphere is an immensely useful tool to the atmospheric scientist. The space shuttle orbiter is a largely untapped resource of extremely high resolution imagery and is available using existing technology.

The potential surface resolution (also called resolving power) of a space-borne imaging system, R_{su}, can be estimated by

$$R_{ru} = \left(\frac{H}{F}\right) R_{su},\tag{D.1}$$

where H is the altitude of the system, F is the lens focal length, and R_{su} is the system resolution (Jensen, 1968). Scene contrast, image motion, and atmospheric conditions can also affect surface resolution, but equation (D.1) is a useful measure of system capabilities.

In the case of the space shuttle the imaging system resolution, R_sy, is given by

$$\frac{1}{(R_{su})^2} = \frac{1}{(R_f)^2} + \frac{1}{(R_l)^2} + \frac{1}{(R_w)^2},\tag{D.2}$$

where R_f is the resolution of the particular film, R_l is resolution of the lens, and R_w is the resolution possible through an intervening window, if one is present (Cox, 1974). In virtually every case, film resolution is the worst of the three and can be used make a rough estimate of the resolution of the entire imaging system.

Resolution of films and windows is governed by the materials used in their manufacture. Lens resolution depends on what is known as the Rayleigh criterion as well. The Rayleigh criterion states that angular resolution is directly proportional to wavelength and inversely proportional to aperature diameter. Consequently, maximum spatial resolution (minimum angular resolution) is achieved at short wavelengths and large aperatures.

Resolution of each of these three components is commonly judged by the accuracy with which they can reproduce the closely spaced parallel lines of a test chart. Film resolutions were listed in Table 2.2. Shuttle window resolution exceeds 700 lines per millimeter and camera lens resolutions each exceed 800 lines per millimeter.

Using information cited here and in Tables 2.1 and 2.2 of this work, it is a simple matter to calcualte the potential surface resolution of the hand-held and large format camera systems. Substitution into (D.1) using values for image 1 (a hand-held camera image) results in a potential surface resolution of .0104 kilometers or 10 meters! Substitution into (D.1) using values for image 2 (a large format camera image) results in a potential surface resolution of .0073 kilometers or 7 meters!

Considering the lower limit of shuttle orbital operations (about 185 kilometers) and the availability of high resolution films and telephoto lenses, even better potential surface resolution is possible.

The narrow field-of-view of long focal length lenses (e.g., 19.5 degrees for a 105 millimeter lens mounted on a 35 millimeter camera) need not be an overriding concern. A potential imaging system's components can be varied according to

$$FOV = 2\left\{arctan\left[\frac{.5(W_f)}{F}\right]\right\}, \qquad (D.3)$$

where FOV is field-of-view, W_f is film width, and F is the lens focal length (Jensen, 1968), to achieve the desired angular coverage.

Using equations (D.1) and (D.3), shuttle-borne imaging systems could be tailored to almost any need. (Imagery need not be restricted to that collected at visible wavelengths.)

END

11-87

DTIC